DOCUMENT RESUME

ED 173 432 TM 009 569

TITLE	Proceedings of the Invitational Conference on Testing Problems (New York, New York, November 3, 1951).
INSTITUTION	Educational Testing Service, Princeton, N.J.
PUB DATE	3 Nov 51
NOTE	115p.
EDRS PRICE	MF01/PC05 Plus Postage.
DESCRIPTORS	Adults; *Cognitive Ability; Educational Testing; General Education; Higher Education; *Human Resources; Individual Characteristics; Military Personnel; Occupational Tests; *Personality Tests; *Personnel Selection; Psychological Testing; *Talent Identification; Test Construction; *Testing Problems

ABSTRACT

Three topics were discussed at this conference: (1) the supply and identification of high level talent; (2) problems of evaluation in general education; and (3) the development of tests for the measurement of non-intellectual functions. Tony Oxtoby and John T. Dailey discussed military and civilian personnel research on talent identification and talent supply. Robert J. Havighurst discussed psychological and sociological factors affecting the supply of talent, and Robert L. Thorndike spoke of future trends. Paul L. Dressel addressed evaluation problems related to general education, including the testing of critical thinking. Major General Lewis B. Hershey, Director of the Selective Service presented the luncheon address. Testing of non-intellectual functions was described by Donald W. MacKinnon--who discussed the measurement of personal effectivness--and Raymond B. Cattell--who discussed the structure and assessment of personality. John Dollard and Silvan S. Tomkins offered further commentaries. (GDC)

INVITATIONAL CONFERENCE ON TESTING PROBLEMS

November 3, 1951

HENRY S. DYER, *Chairman*

¶ The Supply and Identification of High Level Talent.

¶ Problems of Evaluation in General Education.

¶ The Development of Useful Tests for the Measurement of Non-Intellectual Functions.

EDUCATIONAL TESTING SERVICE

PRINCETON, NEW JERSEY LOS ANGELES, CALIFORNIA

FOREWORD

YEARLY, we at Educational Testing Service are impressed by the increasing numbers of those interested in testing who come to the Invitational Conference on Testing Problems. More than 350 persons, the largest attendance in its 15 year history, assembled in New York for the 1951 Conference. This is eloquent testimonial, I think, to the need for and stimulation of the occasion. The published record of the proceedings will, I hope, spread more widely the information and wisdom contained in the papers presented by the distinguished speakers.

This year we have included in the proceedings a transcript of General Hershey's remarks at the Conference luncheon, an innovation at the 1951 Conference.

As is always the case, the Chairman of the Conference has the burden, many months in advance, of insuring an interesting, challenging, and successful meeting. In the capable custody of Henry S. Dyer, I knew, long before November, the Invitational Conference on Testing Problems would meet these goals. I can only add to the many favorable comments and letters already received my own deep and sincere appreciation.

HENRY CHAUNCEY, *President*
Educational Testing Service

PREFACE

The 1951 Invitational Conference on Testing Problems was held, under the sponsorship of the Educational Testing Service, at the Roosevelt Hotel, New York City, November 3, 1951. It was attended by 350 individuals from 31 states and 3 foreign countries. The program consisted of four units: (1) a panel discussing "The Supply and Identification of High Level Talent" (2) an address by Dr. Paul L. Dressel on "The Problems of Evaluation in General Education" (3) a luncheon address by Major General Lewis B. Hershey on "Military Manpower Problems" (4) a panel discussing "The Development of Useful Tests of Non-Intellectual Functions."

These subjects were selected for discussion in the belief that they represent points of contact between test development and some of the most critical problems facing American society today. The pressure on skilled manpower occasioned by the program of National Defense, the widespread movement toward some form of general education in the colleges, and the need for better understanding of personal adjustment in a world of tension—all of these matters generate problems that test makers are being called upon to solve.

The papers and discussions that constitute the main body of this volume demonstrate that there are no ready-made solutions. One discovers, indeed, that there is not even any universal agreement as to the underlying postulates on which fruitful approaches to the problems might rest. The 1951 Conference, like many that have preceded it, has brought these issues into the open and helped to clarify them. The whole discussion, in fact, seems to reflect an increasing concern for finding that series of concepts from which empirical studies can most profitably take off.

The foregoing, however, should not carry the implication that the participants in the Conference are "mere theoreticians." Quite the contrary. Every one of them is actively engaged in research or other activities involving the practical application of psychometric techniques to concrete situations. And for every one of them participation in the Conference involved a sacrifice of valuable time out of a busy schedule. I am happy to have this opportunity to thank them

for the high degree of thoughtful consideration that they gave to the matters with which the Conference dealt, and I feel sure that the reader will find their papers and the ensuing discussions rewarding.

HENRY S. DYER, *Chairman*
1951 Conference

CONTENTS

PANEL I

The Supply and Identification of High Level Talent

The Supply and Identification of High Level Talent

TOBY OXTOBY

THREE BASIC PROBLEMS

THE SUPPLY and identification of high level talent is a major problem to which the Commission on Human Resources and Advance Training has addressed itself for the past year. It is not a problem which can be solved in one year, or in ten years for that matter, and the Commission has only begun to make a preliminary assessment of it. I think that it will be satisfied if it can make only a good start toward bringing the problem into focus so that reasonable solutions of its various aspects may eventually be found.

Different things are often meant by the expression "high level talent." The Commission on Human Resources thinks of it as a person who is able to work and earn a living in the sciences, social sciences, humanities, or professions. It may seem questionable whether members of all of these various fields should be included under the general description of top level talent. It is immediately obvious, however, that they all have one thing in common: a certain minimum level of intelligence is required as well as a fairly long period of formal education in an institution of higher learning. Occasionally, differentiation is made between the fields on the basis that some are more "important" to our nation than others. Such a difference is difficult to defend. All fields are needed as integral parts of our economy; loss of any one field as an active group would have repercussions not only for that field but for many others as well. Again it might be supposed that some fields could be rated more top level than others on the basis that a higher intelligence level is required, but although differences are found among the average intelligence levels of different fields, these differences are slight in comparison with the range of scores within a field and with the over-all difference between these fields and the general population.

Let me again make clear that the definition of high level talent I am using refers to persons who can successfully earn their livelihood as members of these various fields. It is not necessarily concerned with all persons who can earn degrees in these specialized fields. Every year, for example, seven or eight thousand persons receive bachelor's degrees in psychology. Yet a maximum of only 1,500 of these eventually work as psychologists. This example makes clear that in this area our interest is not in identifying persons who will do well in the study of psychol-

ogy, but rather in identifying persons who can end up by actually working in the field.

The initial problem therefore, is to obtain information on what persons now working in the top level areas are like. Studies on the intelligence, academic achievement, motivation, personality, etc. of persons now actually engaged in these fields have not been very common. For this reason, the Commission on Human Resources and Advanced Training is spending a considerable amount of its research time in collecting data on this subject. Studies to date indicate that it is not always the most intelligent, nor the individual with the greatest academic promise who ends up as a member of these fields. An overwhelming number of such persons are, however, quite high on at least one of these variables.

An understanding of the supply of top level talent depends on an understanding of the characteristics of persons now working in the sciences, social sciences, humanities, and professions. Once we know the pattern of intelligence, academic achievement, personality, etc. of persons currently in these areas, we are in a position to determine how many individuals there are in the population who are similar to them. Unfortunately, the number determined by this method is not really the supply of top level talent. It is, however, a sort of least upper bound to the supply.

Since one characteristic of persons now working in these specialized occupations is that nearly all graduated from college, we will find that the younger the group of persons in the population we consider, the larger will be the supply of top level talent. It is clear that college graduates over the next four years are limited to persons who have already entered college, and we know that the graduating class of 1955 will represent a substantial decrease from the number of persons who began college this fall. Likewise, the number of persons who graduate in 1956 is limited by the number who will graduate from high school this coming June. But at this level we find a substantial increase in the supply of high level talent. Of the top quarter of high school graduates in scholastic standing, for example, less than half enter college as full time students. While still in high school this group may be considered as good potential material but those of the group who fail to enter college are lost forever as prospects for our top level fields.

The whole story of why these potentially able persons drop out so early is still not completely known. Perhaps the most common reason advanced relates to the financial status of the student's family. Many studies have shown that the greatest drop out of high level talent is among persons who came from the lower economic groups. Finances themselves are not the whole story, however, since these economic groups are more likely to consider education a waste of time. If the reason for non-attend-

ance is financial, the number of persons continuing on in school could undoubtedly be increased by a more liberal scholarship program. Where the reason relates more to family attitude, however, it would be a much more difficult job to increase the number attending college.

A factor not often emphasized here is that there are other demands made on this group of high level talent. High intelligence has other uses than in the fields requiring higher education. It has often been said that a smart person can be pretty successful in almost anything he tries. Certainly the ultimate location of top level talent depends very largely on the interests and motivations the young person develops either in his home, school, or community. Here there is the two step problem of first finding out whether such individuals are better off going into the specialized fields, and secondly, if an affirmative answer is given, of developing interests in that direction.

This brief sketch of some of the speculations in this area demonstrates that any assessment of the top level talent must depend on a fairly accurate understanding of the reasons why able students drop out at various stages in the educational process. We must also have some knowledge of the motivation, interests, and plans of these groups of top level talent. For this reason the Commission on Human Resources is conducting or cooperating in a rather extensive series of studies of the motivations, interests, and attrition of students at both the high school and college level. This information will give us a reasonably accurate estimate of the number of persons who could under different sorts of inducements be expected to go into the top level fields.

Any consideration of the methods by which young people might be induced to go into these fields must take into account the question of demand for top level talent. Because the demand in most of the specialized fields has always been small in relationship to the total number of persons graduating from college the drop-out of members of the top talent group has in the past been considered more of a loss to the individual concerned than to the nation itself. In the past few years however, there have been indications that like so many other of our natural resources that have come close to exhaustion by a tremendously growing demand, we are beginning to run short on our human resources. A quick glance at the field of engineering should convince you, for example, that the job opportunities and training facilities greatly exceed the number of persons with the capability, the motivation, and the preparation for engineering.

The question of future demands in any field is a difficult one, and it is particularly difficult in fields which require top level talent. The Commission has, however, set for itself as a major task the making of such predictions. These predictions must have as a base growth trends in the fields

for the past twenty or thirty years. Reliable growth curves are not generally available, but by digging up scattered facts from many sources, predicting trends for entire groups on the basis of small samples, and taking advantage of consistencies of certain trends and relationships among the various fields, trend curves are gradually being assembled. From these major trend lines predictions of what is likely to happen during the next ten or twenty years are being made. In addition to trends of what is expected to happen on the basis of what has happened in the past, extra demands for persons in times of emergency are also being estimated. In trying to peer into the future the Commission has been interested in tracing realistic demands rather than setting up arbitrary, perhaps unrealistic suggestions as to "What ought to happen."

In presenting this rather brief outline of the things the Commission on Human Resources and Advanced Training is doing, I have tried to emphasize the close interrelationship of the status quo, supply, and future demand aspects of the problem. In the first place, to identify top level talent we must become acquainted with the characteristics of persons now working in the fields requiring higher education. In the second place we must know how many persons in the population have these characteristics, and the reasons why some of them go into fields in which we are interested while others do not. Thirdly, we must evaluate the future demands so that we can determine how many of the potentially able persons should be trained and can reasonably be expected to be trained.

These problems are so interrelated that they seem to us to be a single problem. And that is precisely the reason why we have undertaken to work on all three at the same time.

The Supply and Identification of High Level Talent

JOHN T. DAILEY

STRETCHING THE SUPPLY OF HIGH LEVEL TALENT

IT IS BECOMING increasingly apparent in these days of world crisis that one of our national resources in shortest supply is the supply of high level talent. It would seem that the supply of such high level talent is grossly inadequate to fill the mobilization demands placed upon it by industry and the defense agencies and still meet the needs of our society for the civilian leadership necessary for its continuance as a free and dynamic way of life. However, it is not necessary to accept this defeatist point of view if we realize that the supply of high level talent need not be regarded as a static supply. As with many other national resources it should be possible through proper research and management not only to increase the efficiency of utilization of the supply presently available but also to increase the total supply. One of the primary missions of personnel research today should be the accomplishment of this objective.

The supply of high level talent will be in part a function of the number of relatively independent dimensions of individual differences that are generally accepted as significant talents. For example, we might define "high level talent" as being in the top X per cent of the general population in any dimension or set of dimensions of individual differences that are generally accepted as being important in human affairs. By this definition the supply of high level talent could range from X per cent to near 100 per cent of the general population depending upon the number and intercorrelations of the accepted dimensions of talent. It can be seen from the above that the so-called "supply" of high level talent may be increased by any one of the following procedures:

A. Increasing the precision with which we can identify and measure differential dimensions of relevant individual differences.

B. Increasing the number of independent dimensions of talent by re-defining tasks or jobs in such a way as to minimize the number of relatively independent dimensions of individual differences involved in success in each task or job.

C. Increasing the efficiency of training in order to minimize the magnitude of the "critical amount" of talent necessary for success in each task or job.

D. Insuring that each individual in the population has a full and adequate opportunity to develop and use whatever latent talent and poten-

tial for development that he may possess.

Much of the military and industrial personnel research of the past decade has been oriented toward one of the first three objectives above. It is felt that the fourth objective, that of insuring that each individual in the population has a full and adequate opportunity to develop and use whatever latent talent and potential for development that he may possess, falls more properly within the domain of the educators and social scientists. However, it is believed that the personnel psychologists should realize that the achievement of this fourth objective might do more in the long run than their own efforts to stretch the supply of high level talent. Accordingly they should lend every possible assistance and encouragement to the research efforts of those active in this field.

Most of the military research oriented toward stretching the supply of high level talent falls within the areas of personnel selection and classification, job and qualifications analysis, training research and human engineering. Relevant aspects of this research will be commented on later under these four categories.

Let us now return to our original definition of talent as being in the top X per cent of the general population in any dimension or set of dimensions of individual differences that are generally accepted as important in human affairs. We can see that if talent is regarded as unidimensional (general intelligence for example) the supply of high level talent will be X per cent. However, if talent is regarded as existing in two independent dimensions, then the supply of high level talent becomes X per cent plus $\frac{X(100-X)}{100}$ per cent or very nearly 2X per cent if X is small as would usually be true. Of course the increase becomes considerably less if the two dimensions are correlated. The supply will be asymptotic to 100 per cent as the number of independent dimensions becomes large, and even with moderately correlated dimensions the supply can become relatively large with a reasonably small number of dimensions.

Here it should be pointed out that the supply of talent is related both to the number and intercorrelation of dimensions of relevant individual differences in individuals and also to the number and intercorrelations of these dimensions involved in each job or task. For example, if talent be regarded as existing in two dimensions (Let us say general intellectual aptitude and physical status) then the supply of talent available to fill the existing high level jobs could be as low as X per cent multiplied by X per cent if each such job or task required that the incumbent be in the upper X per cent in each of the two dimensions of talent. This is not a mere hypothetical illustration since the armed forces frequently require both high physical and intellectual status for many assignments.

The reasons for this are obvious and are usually justifiable. However, it should be realized that this seriously reduces the supply of high level talent available to the armed forces and should be avoided except where absolutely necessary.

It would be well at this point to look into the question of the probable number of independent dimensions of talent. While the number of independent dimensions of talent is probably much smaller than the number of orthogonal factors possible to isolate and rotate, there is good evidence that talent does exist in at least several fairly independent dimensions. Among these appear to be: Intellectual aptitude, mechanical aptitude, physical condition, perceptual-motor aptitude, motivation and interest factors, and temperament factors. Each of these areas may, of course, be resolved into numerous sub-dimensions, many of which may be relatively independent. Considerable success has been realized in developing comprehensive classification test batteries consisting of tests representative of each of the above six areas. Such batteries are now widely utilized in the various armed forces and by industry. To the extent to which they yield differential measures of the various dimensions of talent they increase the available supply of high level talent as contrasted to the older days when talent was identified largely in terms of general intelligence. It should be pointed out that the efficiency of most existing classification test batteries should not be overestimated as most of their discrimination is in the intellectual and mechanical areas. Generally the agencies utilizing them are well aware of this and are concentrating a large portion of their research effort upon improving the coverage of their batteries in the other areas.

The classification test batteries yield information making it possible to identify the high level talent with greater precision and to increase the available supply. The problem still remains of achieving optimum efficiency of utilization of the supply identified. This is the so-called "classification" problem as recently defined by Thorndike. The solution to this problem consists of selecting the man-job assignment configuration that yields the highest productivity sum for the group involved. Considerable work on this problem has been accomplished by the armed forces and several possible mathematical solutions have been developed and are being tried out experimentally in empirical studies. They show promise of aiding the armed forces to achieve optimum efficiency of utilization of their limited supply of high level talent.

Recently the armed forces have completed several interesting studies of the intercorrelations between criteria of training success in various job areas. It has been found that such criteria as technical type school grades for different jobs often show a disturbing amount of common

variance. It is to be doubted that this always reflects a true overlap in the talents involved in actual job productivity in the different areas. There is good reason to believe that much of this overlap may be an artifact of the grading systems which include many written achievement tests and of the curricula which involve a considerable amount of verbal and written subject matter. Classification test batteries are usually validated against these school grades and are thus severely limited in the amount of discrimination that they can achieve between such criteria. The three services are acutely aware of this limitation and are making strenuous efforts to develop usable criteria of on-the-job effectiveness.

This type of spurious overlap between school grades may cause inefficient usage of talent in still another way if there is an appreciable amount of attrition on the basis of school success. This will be true to the extent that the spurious overlap in such factors as verbal facililty for example are not involved in the actual performance on the job.

In addition to the problem of stretching the supply of high level talent the nation also faces the equally acute problem of stretching its supply of low level talent. In this connection, low level talent may be regarded as those cases where X per cent of the general population represents a relatively high percentage but one that is definitely less than 100 per cent. Examples of such talent might be artisans, non-technically trained enlisted men, etc. It is not believed that the problems regarding low level talent are qualitatively different from those regarding high level talent. The same basic principles for stretching the supply of high level talent should also apply to stretching the supply of low level talent. If anything, the results of the application of these principles with low level talent should be even more pronounced than for high level talent since the low level jobs are usually less apt to require a number of separate talents.

Perhaps the most promising method of stretching the supply of high level talent is by increasing the number of independent dimensions of talent by redefining tasks or jobs in such a way as to minimize the number of independent dimensions of individual differences involved in success in each task or job. This is often referred to as work or job simplification. An eminently successful example of this is the platoon system in football. Not only do the players usually specialize in offensive or defensive play, many also specialize to the extent, for example, of only holding the ball for extra point attempts. This has undoubtedly raised appreciably the average level of skill displayed in each of the elements of modern football play. Under the old system of relatively non-specialized players the supply of high level football talent was much more limited than today. In those days a boy might have considerable talent in

catching passes, for example, but be too light to play defensively. Another might have considerable talent for defensive line play but be relatively weak at offensive play, perhaps because of lack of speed. Still another might be very talented as a punter but be weak at all-around play either offensively or defensively. By means of job simplification the supply of high level football talent was apparently greatly increased.

This matter of job specialization is of urgent importance to the military and lies at the heart of many of their personnel problems. However, the answers to these problems are not easily achieved. The nature of the military situation is often not very appropriate for the type of free substitution that permits football or industry to go so far in the direction of specialization. In combat there is no time out between plays and often the space on a ship or in an aircraft severely limits the size of the crew. Thus each crew member must be responsible for a variety of types of duties and also be able to take over many duties of other crew members in emergencies. In addition a high degree of specialization creates tremendous administrative problems where personnel frequently have to be transferred from one organization to another. However, because of concern over the limited supply of both high and low level talent, the services are carrying out research on job simplification in order to permit the maximum use of specialists under the limitations within which they must operate.

Closely related to the work on job simplification is the work on human engineering. In this area one of the basic problems is to determine how to design equipment etc. in such a way as to minimize the complexity of the task of operating and maintaining it. To the extent to which they accomplish this purpose they minimize the number of relatively independent dimensions of individual differences involved in success in the tasks or jobs based on the equipment and thus aid in stretching the supply of high level talent.

As pointed out earlier, another way of stretching the supply of high level talent is to increase the efficiency of training in order to minimize the magnitude of the "critical amount" of talent necessary for success in each task or job. The armed forces are highly aware of this and are investing heavily in research aimed at increasing the efficiency of training. A large amount of research is now underway in determining the optimal methods of utilization of such training aids as synthetic trainers, films, mock-ups, and even television. Further work is directed toward the area of training methods, improving the quality of instructors, etc. The services are even sponsoring considerable fundamental investigation into the nature of the learning process itself especially in such areas as motor and perceptual learning. All of this research should

assist greatly in stretching the supply of high level talent.

Let us now look at a case history of a hypothetical high level job for which talent is in short supply and see how we might apply the above principles to stretch the supply of talent for the job. We will suppose that the job consists of maintaining equipment Y which is very intricate and difficult to maintain. Typically we might encounter the situation where the selection standards for entry into training for this job are quite high and yet there is still a high elimination rate during the training course. It is found that if the present selection standards are raised it will not be possible to fill the training quotas and still meet other equally important training quotas requiring similar talent. What is needed in this situation are procedures which will reduce the elimination rate during training without decreasing the supply of talent available for other commitments and without reducing the efficiency of the graduates on-the-job. It is believed that any one or combination of the following procedures could help appreciably to accomplish this objective and thus stretch the supply of high level talent.

1. If the differential validity of the classification procedures is increased it will be possible to lower the training elimination rate without decreasing the supply of talent remaining for assignment to other high level jobs. It should be pointed out that this cannot be accomplished merely by raising the correlation between the selection standards and training success if this is accomplished at the expense of increasing disproportionately the correlation between the selection standards for the various jobs competing for the same manpower.

2. The equipment and maintenance procedures could possibly be re-designed to simplify maintenance problems without loss of operational efficiency.

3. Improved training devices and procedures could make it possible to graduate a higher proportion of students without raising entry standards or decreasing on-the-job proficiency.

4. The training curriculum could be revised to eliminate all elements not directly contributing to on-the-job proficiency. In some cases it might be possible to decrease the emphasis in theory and mathematics, for example.

5. It might be possible to redefine the job and divide the duties into two separate jobs each requiring a lower level of talent than the original job, or it might be possible to redefine the job in order to have one high level man with two more specialized and lower level assistants do the jobs originally set up for three high level men.

6. Finally by better management and administrative procedures we may decrease the total personnel at all levels necessary to keep the equip-

ment maintained thus decreasing the drain on the manpower supply.

In conclusion, it may be reiterated that the supply of high level talent is not a static supply but a supply that apparently can be stretched considerably as the result of research now underway in both military and civilian personnel research programs.

The Supply and Identification of High Level Talent

ROBERT [redacted] HAVIGHURST

SOCIOLOGICAL AND PSYCHOLOGICAL FACTORS AFFECTING THE SUPPLY OF TALENT

BY TALENT, let us mean very high ability in some socially valuable area. There are many socially valuable areas, including such diverse ones as abstract thinking, tennis playing, leadership, and salesmanship. But for the sake of this discussion let us limit ourselves to intellectual and artistic areas.

How high must ability be when we call it talent? For the purpose of this discussion let us consider two levels of ability. By *talent* let us mean the upper two per cent, and by *high ability* the upper twenty per cent.

SOCIOLOGICAL FACTORS AFFECTING THE SUPPLY OF TALENT

In order to discuss the supply of talent we must make some assumptions concerning the visibility of talent. We may assume it is all in sight and that we can recognize it, or we may assume that a considerable part of the supply of talent is not readily visible to us.

The Distribution of Talent Among Social Groups

Confining our attention at first to talent which is visible we may inquire into its distribution among various social groups. First let us consider the kind of talent which is indicated by high verbal intelligence as measured by the ordinary intelligence tests. This is distributed over the entire socio-economic range, but there is relatively more of it in the upper reaches. Of the top fifth of the population, in verbal intelligence, approximately half come from the upper third on the socio-economic scale, while the other half comes from the lower two-thirds, which is often called the working class.

It is not certain whether there is an even greater segregation of very high verbal intelligence by socio-economic status. Terman, in his Genetic Studies of Genius, which dealt with children in the upper one or two per cent of ability, found a very high proportion of them to come from the higher socio-economic levels, but his procedure for selecting talent may have overlooked some at the lower socio-economic levels.

Turning to musical and artistic talent, the distribution does not seem to be related directly to socio-economic status. Musical talent appears to be related to cultural background—the Italians specialize in grand opera, Jews in pianists and violinists, Negroes in singers, and American born composers in jazz. Jews have been unusually productive of all kinds of artistic talent, as well as of high verbal intelligence.

There are some sex differences in

the distribution of artistic talent, with men showing more very high level records. It is interesting to note that the great musical composers have been men, almost without exception. Men have provided most of the great painters and poets, but women have come much closer to men as great actors, singers, and novelists.

The known distribution of visible high level intellectual and artistic talent can be explained on the theory that each social group possesses cultural habits and attitudes that tend to bring out certain kinds of talent and to suppress others.

What about talent that does not come to light? There are two possible reasons for a talented person going undiscovered. One reason is that the tests or other means used to discover talent are inadequate. Intelligence tests certainly fall somewhat short of measuring intelligence equally well in all socio-economic groups. Allison Davis has found with his new "culture-fair" test of intelligence that some slum-dwelling youngsters who appear to have about average intelligence on the usual tests actually show a very high level of intelligence on the new tests. We have no good way of estimating the amount of unrecognized talent due to inadequate tests. Certainly a great deal of artistic and musical talent goes unrecognized because the tests of musical and artistic aptitude are inadequate.

The second reason for talent not coming to light is that it may be suppressed—not permitted to develop—by a poor environment. The geneticists tell us now that every human characteristic which has a hereditary basis is influenced in its actual appearance by the environment. Talent probably needs two kinds of environmental encouragement for it to develop. First, the individual needs food and physical care and character training as a child so as to grow up with normal physical energy and enough self control and self assurance to be a purposeful, independent human being. Second, the individual needs a social environment which teaches him to exercise his talent and rewards him for it.

Probably a relatively great amount of talent is suppressed by environment that fails to stimulate and reward its development, especially in the economically underprivileged groups such as American Negroes and Mexicans.

Thus, if we accept the idea that talent or the potentiality of talent exists without being observed, the writer believes that the known facts about talent are best explained by a theory that the potentiality for talent is inherited with little or no group differences among the races and social groups of mankind, and that the appearance and development of talent in individuals is a result of environmental stimulation and guidance. If this theory is accepted, then any social or racial group which does not display many talented individuals may be presumed to provide a social environment which does not stimu-

late the development of talent. This is probably the reason that women seem to show less talent than men in America.

Psychological Factors Affecting the Supply of Talent

The principal factor within the individual which affects the supply of talent is motivation. Motivation is necessary to the development of intellectual or artistic talent—the motivation being a desire to seek training and a willingness to sacrifice other desires while undergoing training. Lack of motivation appears to be the principal reason why visible talent is not trained and developed in this country. Lack of motivation appears to be more powerful than lack of money in reducing the supply of talent.

To explore the motivation factor further, we may concentrate our attention on the question of going to college. Talented youth who do not go to college are not likely to get the training that will enable them to develop their capacities to the point where they will be of most use, socially. Most youth with visible talent finish high school. But a considerable number of them do not go on with formal training after high school. Hence it is well to focus attention at this critical point.

*The Probability of Going to College**

The probability that a given boy or girl will go to college depends on the following factors: mental ability; social expectation, or what the family and the society expects of him; individual motivation, or what his own life goals are; financial ability, in relation to the cost of continued education; propinquity to an educational institution. This proposition can be put into the form of a mathematical equation:

$$p = a \times \text{mental ability} + b \times \text{social expectation} + c \times \text{individual motivation} + d \times \text{financial ability} + e \times \text{propinquity},$$

where p is the probability that John Doe or Ruth Roe will go on to a post high school institution of learning; and a, b, c, d, and e are constants to be determined by empirical studies of boys and girls in the United States.

Mental Ability.—A half-dozen recent studies indicate that 40 to 45 per cent of boys and girls in the upper fifth of the population as measured by ordinary intelligence tests go to college. The actual proportion varies from one type of community to another, but remains between about 35 and 50 per cent.

As we go up the scale of intellectual ability we find higher proportions of youth entering college. Phearman found that 92 per cent of the top 2 per cent of Iowa high school seniors (measured by the Iowa tests of Edu-

* The following pages are taken from a memorandum prepared by the writer for the Commission on the Financing of Higher Education. The complete memorandum is to be published as an appendix to *Who Should Go To College?* by Byron S. Hollinshead, and published by the Commission.

cational Development) went on to college. Terman found that 90 per cent of his gifted group entered college and 70 per cent graduated. Goetsch, with data from the depression decade, found that 63 per cent of Milwaukee high school seniors with IQ 140 and over entered college, while the proportion dropped to 22 per cent of those with IQ in the 117-119 range. The Phearman and Terman figures appear to the writer to suggest too high an estimate of intellectually talented youth going to college. Phearman's and Goetsch's figures must be reduced by the group (a small one) of talented youth who did not finish high school. Terman's group probably had an unduly high representation of children of higher socio-economic status, who would be more likely to go to college because of social expectations.

It is probable, however, that at least 75 per cent of youth of very high intellectual ability enter college, and 60 to 70 per cent of such youth finish a four-year college course.

Social Expectation or Social Motivation.—If a child grows up in a family which expects him to go to college, he is very likely to do so. Since family attitudes toward education differ systematically from one social group to another, we can predict that young people from one social group will very probably go to college, while young people from another group will be very unlikely to go to college.

The social groups most likely to send their children to college are the urban business and profesional people. Those least likely to expect their children to go to college are the urban working-class group and the small-scale farmers.

The most useful social classification of people in terms of their attitudes toward higher education is the socio-economic classification of Warner. A simplified version of this is the following.

a. *Upper and Upper Middle Class.*—Providing about ten per cent of the children, this group sends about 80 per cent of its children to college. Those who do not go to college fall into two groups: (1) the daughters of upper class families who terminate their formal education with a "finishing school" of high school level and (2) the children whose mental ability is too low to permit their success in any kind of post high school institution.

b. *Lower Middle Class.*—This group consists of white collar city workers, people in semi-professional occupations, small business owners, highly skilled artisans, foremen, and owners of "good" farms. They produce some 30 per cent of the children. Approximately 25 per cent of these boys and girls go on with post high school education. Thus it is the exception rather than the rule for a lower middle class youth to go to college. The expected thing for him is to graduate from high school and go to work. No doubt more boys and girls of this class would go to college if they had

scholarship aid, for the economic means of a typical lower middle class family would not permit sending a child away to college. Nevertheless, financial need is not the sole barrier to college-going; there is also a considerable barrier of social expectation and a barrier of individual motivation.

c. *Working Class.*—This group—the upper lower and lower lower classes in Warner's terminology—consists of unskilled, semi-skilled, and some skilled urban workers, most farm tenants and all share croppers. Sixty per cent of the children in the United States come from this group. Less than five per cent of these children go beyond high school in their education. It is even unusual for boys and girls of working-class families to finish high school. The majority of them drop out of school at the legal school-leaving age. There is a strong social expectation in the working class that a boy or girl should quit school and start to work and earn money as soon as he is able to do so. The idea of going to college is not a popular one in working class circles. A boy or girl must surmount a barrier in the social expectations of his friends and neighbors and his family if he goes on beyond high school.

Other Classifications.—There are several other ways of classifying people with regard to social expectation of going to college. Color is one. The work of the New York State Temporary Commission on Need for a State University showed that only one-eighth as large a percentage of Negro as of white youth of college age were enrolled in higher education. This, however, is largely due to socio-economic differences between Negroes and whites.

Ethnicity is another social factor of some consequence. In Connecticut 49 per cent of high school graduates of Italian origin applied for college compared with 57 per cent of Catholics, 63 per cent of Protestants, and 87 per cent of Jews.* Here it seems that Italian ethnic background reduces probability of going to college, while Jewish background (if we may call this ethnicity rather than religion) favors going to college. Here, too, the socio-economic factor was certainly a large element.

Urban residence implies a slightly greater social expectancy of going to college than rural residence. There is a slight and probably decreasing tendency for farm families to be less favorable than city families of the same socio-economic level toward higher education for their children. However, the principal factors reducing college attendance of rural children are socio-economic and distance factors.

Individual Motivation. — Whatever the social expectations that press on him from without, the individual has within himself a set of purposes which may or may not include a desire for

* Stetler, Henry G. "College Admission Practices with Respect to Race, Religion, and National Origin of Connecticut High School Graduates." Hartford, Connecticut, State Interracial Commission, 1949.

higher education. If he has such a desire, and a strong one, he will surmount social and economic barriers to get into college; if he has little individual motivation, he will not get into college unless he is pushed strongly.

For example, a boy whose father was a janitor, whose mother and father had not gone beyond the eighth grade in school, was so imbued with the ambition for higher education that he wrote, at the age of 16, "The worst thing that could happen to me is that I should fail in my plans to go to college and enter into a profession." He secured a scholarship and went on to college. On the other hand, a boy whose parents had a small business and wanted their only son to go to an engineering school, refused to do so and took a job as clerk in a store in his home town, saying that all his friends were in town, his girl friend was there, and he liked working in a store. So why should he go off to college? Both boys were intellectually superior but the factor of individual motivation determined which one should go to college.

Carson McGuire in his study of adolescent social mobility* identifies three groups that have the individual motivation to go on to post high school education. They are:

1. *The high status static.*—This is a person of upper or upper middle socio-economic level, who has the typical educational attitudes of his social group. Though not upward mobile, he will go on to college because that is normal for his group.
2. *The climber.*—This is a lower-middle or a working class youth who has a solid and realistic ambition to "get ahead" in life. This person has friends among boys and girls of higher social status, spends time in their homes, and absorbs their educational attitudes. He has a good mind, a strong personality, understands that self-control and hard work will be required of him, and is prepared internally to make the sacrifices necessary for the achievement of higher education and consequently social mobility.
3. *The strainer.*—This is a lower middle or working class youth whose goals in life are mixed, and whose own personality is vacillating. He wants to "make good," yet is not completely sure within himself what this means. He makes friends with boys and girls higher on the social ladder, but is never quite sure that he desires their way of life. In the end it is something external to him, such as a GI Bill of Rights, that decides whether he shall go on with his education.

Financial Ability.—To a considerable extent the ability to pay for a college education is systematically related to socio-economic status. The upper and upper middle group can afford to help their children get a

* McGuire, J. Carson. "Adolescent Society and Social Mobility." Unpublished Ph.D. Dissertation, The University of Chicago, 1949.

higher education though sometimes at a considerable sacrifice, especially if the number of children in the family is large. For the lower middle group it is a great financial burden to send even one child away to college, though it is possible to support a son or daughter in a college at home, especially if it is tuition-free. Furthermore, the student is expected to contribute something through working during the summers. For the working-class group there is little possibility of financial help from parents, except where there is only one child, or only one child in the family is going to college. Furthermore, the loss of the earnings of a youth who is going to college instead of working is felt keenly by such a family. Accordingly, most young people in this group who go to college must support themselves almost completely, through employment or through scholarship grants.

Propinquity. — The percentage of youth going to college is usually greater in communities with colleges than in communities without colleges. This is due partly to the financial advantage of being able to live at home while going on with one's education, and also due to the fact that the goal of higher education is more familiar and more attractive when it is visible in the home town. Youth who would not make the effort to go away to school will continue their education if school is brought to them, as it were. The social expectation of going to college is increased in a college town.

Summary of This Theory.—To apply this theory of the determining factors of post high school education, the writer has drawn up a kind of paradigm in the form of a table which combines three of the five factors. The other two factors are kept constant in the case of a particular table. These constant factors are propinquity to an educational institution and mental ability.

TABLE I.

PROBABILITY OF GOING TO A POST HIGH SCHOOL INSTITUTION, FOR YOUTH OF SUPERIOR MENTAL ABILITY

Probability levels are indicated in the respective cells.

Socio-Economic Status	Individual Motivation		
	Low	Medium	High
High (Upper & Upper Middle)	3 Doubtful	2 High	1 Very High
Medium (Lower Middle)	6 Low	5 Doubtful	4 High
Low (Working Class)	9 Very Low	8 Low	7 Doubtful

Table 1 shows the probabilities of going to college for youth of superior mental ability. Cells number 5 and 7 are especially important because they contain relatively large numbers of people. For those in the upper 2 per cent of mental ability all of the probabilities should be increased somewhat over the levels indicated in Table 1. However, even for this group of outstanding intellectual ability there are sure to be individuals in cells 6, 8, and 9 with low individual motivation or low social motivation, who will not go to college.

Application of This Theory. — The writer has applied this theory to two age-groups in a small midwestern city where the necessary data were available. The theory worked quite well to predict the proportions of youth of superior ability who continued their education beyond high school. The following examples illustrate high and low motivation in relation to college-going.

Some Cases of High Motivation. — A was a boy in the 1926 group, cell 7, IQ 122, the son of a widow, his father having been a factory worker. A was a faithful attendant at Sunday School where his teacher, a professional man, took an interest in him and encouraged him to plan for college. His pastor did likewise. They recommended A for a scholarship to a church-related college and their recommendation was supported by the high school faculty who held a high opinion of A. A's mother worked to support herself and could not help him. A went off to college and was inducted into the armed services after about a year of college work. When he was discharged he went back to college on a GI scholarship and later attended a professional school. Although the GI scholarship was a great help to him, there is no reason to doubt that A would have worked his way through college if he had found this necessary.

B was a girl in the 1932 group, cell 4, IQ 143, one of four children in a lower-middle class family. Her father operated an automobile service station. Her mother, though not a college graduate, was much interested in school affairs and was active in church and women's club work. Even though it meant a considerable sacrifice for them, they determined to give B a college education, and they made their decision as early as her freshman year in high school. B was always quite sure that she would go to college. She was encouraged in this ambition by her teachers and her school grades were in the upper tenth of the class, thus giving her some reason to hope for scholarship aid. She went off to college as a matter of course, though not one of her five or six close friends did so.

Some Cases of Medium Motivation. — E was a boy in the 1932 group, cell 5, IQ 139, whose father was a master plumber. Though neither he nor his wife had gone beyond high school, they wanted their only son to go to college. This they urged on E for the last two years of his high school

career. They could afford to pay his expenses. They talked with the high school teachers and asked for their help in persuading E to go on with his education. But E merely shrugged his shoulders. He had a part-time job as clerk in a food store, and he liked this kind of work. Besides, he said, "My friends are all going to stay in town; my girl friend is here; I like this place. Why should I leave it?" At last accounts E was doing very well in his full-time job in the food store.

F was a girl in the 1932 group, cell 5, IQ 134, the most talented person in her class. Her parents had neither of them graduated from high school, but they wanted their oldest daughter to do what was best for her. They had a small business and told her they could help her substantially if she wanted more education. They did not urge her to go to college. For a time she thought of going to a music school for she had enough musical ability to justify encouragement from her high school music teacher. F was always in the lead of things in school, a great organizer, quite popular. She played the role of idea-giver in her age-group. She promised to be a leader wherever she might locate herself. To the surprise of her teachers she took a minor office job at home, making no effort to get into college. Before the end of the year of her high school graduation she was married.

G was a boy in the 1932 group, cell 8, IQ 123, who as early as his tenth year showed remarkable visual imagination on tests, and was diagnosed by a psychologist as a boy who might make an outstanding architect or mechanical engineer—one in a thousand. G's father was a factory hand—neither he nor his wife had graduated from high school. They had one other child and could have helped G substantially with college expenses. G was no more than a high average student in the verbal subjects of the high school curriculum, but in mechanical drawing and shop work he was remarkable. G had two good friends neither of whom planned to go to college. For a time he thought of going to a school which trained draftsmen, but he ended by taking a job in an automobile service station.

The Role of Motivation

For boys and girls in the upper fifth of the population in general intelligence it appears that about 4 out of 10 go to a post high school institution, another 4 out of 10 do not go because they lack sufficient motivation, and 2 out of 10 would go if they had liberal financial assistance. This is the conclusion from a survey of recent studies of the reasons given by the abler high school graduates for not continuing their education. However, most of those who lack motivation also lack financial means, and they would need financial help if their motivation was changed to cause them to wish further training.

For boys and girls in the upper 2 per cent in intelligence there is a

much larger proportion in college—possibly as many as 7 or 8 out of 10.

Conclusions

This treatment of the motivational factor as it affects the supply of talent indicates the great importance of motivation for higher education in determining the supply of "high ability" (upper fifth) people who take college-level training. Apparently the motivation for college training is much greater among the top 2 per cent of the population than it is among the top 20 per cent. However, the data to establish this point exactly are not available. The writer inclines to the view that even in the top 2 per cent there may be as many as a third of the group who do not seek college-level training.

Finally, it should be noted that the studies of *who goes to college* are based primarily on measures of talent which (1) may overlook a good many people of high artistic ability, (2) may overlook a number of people of deviant cultural groups who do not show their real abilities on the tests that were used, and (3) almost certainly omit a number of people whose talent has been suppressed (or not developed) because they grew up in a social environment that did not stimulate and guide them to develop their talent. The existing methods of discovering talent severely underestimate the amount of potential talent in our population.

Yet it is probably true that, due to our system of universal education, to a relatively high degree of economic opportunity, and to our culture which rewards or at least tolerates a wide variety of talents, our society probably has the greatest amount of visible and developed talent that any human society has had. We probably develop more of our potential talent than the Greeks of Fifth Century Athens, the Italians of the Renaissance, France of the Age of Reason, and Britain in her Imperial Nineteenth Century.

The Supply and Identification of High Level Talent

ROBERT L. THORNDIKE

A LOOK AHEAD

Dr. Oxtoby has outlined for you the research program of the Commission on Human Resources and Advanced Training. His group has chosen as its working definition of high level talent "a person who is able to work and earn a living in the sciences, social sciences, humanities, or professions." Because work in these areas tends to require as a prerequisite an extended period of academic training, the definition of talent tends to be formulated in terms of ability to succeed in academic work.

We cannot quarrel with the Commission if it chooses to formulate its task in terms of this definition of talent. This certainly provides ample scope for the activities of an able group of investigators. However, we should recognize that the definition is somewhat restrictive, not being based, for example, on specially high abilities in dealing with things or with people, as distinct from abstract concepts and ideas.

Within the framework set by its definition of talent, the Commission has been interested in determining (1) the number and characteristics of the present group of workers in the various scientific and professional fields, (2) the reasons why persons who apparently have the ability necessary to train for and enter these fields do not do so, and (3) the anticipated future demand for these special groups of individuals. I wish Dr. Oxtoby had had the time this morning to give you some part of the fascinating picture which he and Dr. Wolfle have been piecing together from the accumulated records of past testing and the fragmentary investigations of a number of research workers. Some of the material, such as Dr. Wrenn's study of Iowa and Ohio State Ph.D.'s is available in published form, but much of the picture is new—at least to me—and I have gotten only fragments of it in occasional conversations. This time might better be spent in having them present some of the picture than in having me talk about it.

Dr. Havighurst also restricts himself, for the purpose of his presentation, to intellectual ability, again focussing his attention particularly on the completion of higher education, though he makes a passing obeisance to artistic and other forms of talent. Within the domain of academic talent, he draws attention to the talent which is latent and that which is expressed in actual testable ability. The reality of this difference has been made

abundantly clear by the difference in test performance between U. S. recruits in World War I and World War II. Test performance indicates that well over 80 per cent of World War II soldiers equalled the average recruit of World War I. I think that we would all agree that this represented merely a more complete realization and recognition of academic or test-taking potential inherent in our population rather than a basic change in the poulation itself. It is also a type of confirmation of the conviction expressed by Dr. Havighurst that latent academic talent is made more fully manifest in our country today than at any other time and place.

Dr. Havighurst goes on to discuss the factors which prevent those who show evidence of the ability to undertake college level training from actually undertaking it. Without minimizing the importance of purely financial barriers, he emphasizes the importance of sub-group cultures and individual motivation. These points are perhaps worth speculating on from a slightly different angle.

I seem to sense in, or project into, this whole session this morning, the bias of a college-centered culture. That is, we seem to find it necessary to look for reasons why people do not go the college rather than reasons why they go. Granted that the jobs open to college graduates tend to be more prestige-laden in our society, it is far from unequivocally clear that the financial returns are consistently higher, or that college graduates lead a generally more contented or satisfying life. That is, unless we assume that status, responsibility and work under intellectual pressure are good, and to be desired, we may be more amazed at the numbers who go to college than the numbers who do not. In feeling that more of those who might succeed in college (and beyond) should pursue higher education we may be imposing a very biased set of values upon our countrymen. Our national peril may require that we do so. But we must recognize the culture-centered value judgment which is involved when we imply that it is better to be a harassed high school teacher than a relaxed radio repairman.

Dr. Dailey sets as his definition of talent "being in the top X per cent of the general population in any dimension or set of dimensions of individual differences that is generally accepted as being important in human affairs." This is in sharp contrast with the other two speakers, who have been rather exclusively preoccupied with talent for higher education and a single dimension of general intelligence. I find in Dr. Dailey's presentation, in part by implication, several notions on which I would like to comment. These are (1) that there are many dimensions of human ability, (2) that these can be divided into those that are important in human affairs and those

that are not, and (3) that the socially important dimensions are only slightly related to one another. If these notions are accepted, the number whom it is appropriate to consider "talented" can be very substantially extended.

The number of dimensions which can be conceived of is limited only by the human vocabulary, so the first notion can only be considered in connection with the other two. That a categorical distinction can be made between important and unimportant dimensions of variation is, I imagine, one that few of you would explicitly endorse. In practice we have a hierarchy of valuations of talents which is quite stable and well-structured. Many studies in civilian life have shown the doctor, college professor, and the like to stand high in the prestige hierarchy of rating groups (themsevles typically individuals from the intellectual subculture.) A recent study which we carried out among Air Force officers showed a similarly well-structured hierarchy in the evaluation of Air Force enlisted specialties. The electronic technician is more highly valued than the cook, and the cook in turn is more highly valued than the mimeograph machine operator. By the same token then, we have a hierarchy of talents which grades imperceptibly from those which are deemed (by some particular group) most important in human affairs through those which are judged moderately important down to those which are judged of trivial importance. There is no point at which it is possible to make a distinction between the important and the unimportant, nor are those that are important equally so.

Drs. Oxtoby and Havighurst have chosen to deal with one dimension which is sort of a centroid of a very high priority cluster. Dr. Dailey has emphasized the other dimensions which grade down in importance from that one. My point, if I have one, is that many of the additions which Dr. Dailey will make to the pool of the talented will be additions in less highly esteemed dimensions of talent. Useful as exceptional talent in spool-packing may be, we cannot accept it as an equally-valued extension of our pool of talented surgeons.

We may inquire next how independent the socially important talents will be found to be. Dr. Dailey mentions intellectual aptitude and mechanical aptitude as "fairly independent dimensions" of talent. Let us take the Bennett type of mechanical items as representative of the mechanical domain and tests of word knowledge and arithmetic reasoning as representative of the intellectual domain. What correlation will we find between error-free measures of these attributes in the general adult population? During the past year, we have pushed the doorbells of a good many thousands of households in a number of communities and asked an adult male in each household, where we could find one, to

take brief tests of the three types which I have just indicated. We were primarily interested in survey methodology and in the cooperation which it was possible to elicit by this approach—which was in general quite good. However, we have been carrying out some analysis of the test results, and this would lead me to estimate that very reliable measures of these three abilities would show correlations in the adult male population approximately as follows:

Word knowledge vs. mechanical comprehension, .45

Arithmetic reasoning vs. mechanical comprehension, .65

Word knowledge vs. arithmetic reasoning, .70.

Taking an average of word knowledge and arithmetic reasoning as representing one domain, and mechanical comprehension as representing the other, how much should we expect to augment the supply of talented by considering both dimensions instead of only one? Taking the two degrees of talent which Dr. Havighurst has referred to, the top 2 per cent and the top 20 per cent, I have gone to Mr. Pearson's excellent tables of bivariate normal distribution to get an estimate. I find that I would add an additional 1.5 per cent and 10 per cent at the two levels which I have mentioned. That is, I would then have totals of 3.5 and 30 per cent rather than 2 and 20 per cent respectively.

These increments are not to be sneered at, and the importance of considering additional dimensions in our thinking about talent is very real. However, the additional gains from still more dimensions, if they are correlated with several of those which we already have, will become progressively less. We must not permit analyses based on two or three dimensions or ones which are uncorrelated to cause us to get too starry-eyed about the "classification" point of view.

The typical situation in getting on in any job is that a combination of talents is required. Dr. Dailey has emphasized the gains to be achieved in effective supply of personnel by reducing the numbers of talents involved in a given job. This may be illustrated again by the simple example of the mechanical and the verbal tests. What is the situation if the task requires, as it may, that the person be in the top 2 per cent of the population on *both* the verbal and mechanical tests? Our available supply is then 3 men in a thousand. If we can reduce the verbal requirement to being in the top 20 per cent, we have available another 9 men per thousand, and if we require only that he be average, we qualify still another 6 per thousand.

Multiple requirements are the really crippling problems in our supply of talents, and any steps which we can take to reduce them represent most worthwhile gains.

In conclusion then, while I feel

that we must expand our concept of the talented to include important attributes other than those of abstract intellect, such as abilities in dealing with things and people, and that in so doing we will expand very considerably the total pool whom we will consider talented, the facts of correlation of abilities and multiple requirements by life tasks are such that I doubt that we shall ever be able thereby to achieve the millenium or the late Huey Long's objective of "every man a king."

DISCUSSION

PARTICIPANTS:

ROBERT D. NORTH, HENRY S. DYER, HENRY W. BRAGDON, HAROLD GULLIKSEN, WILLIAM J. E. CRISSEY, JOHN T. DAILY, WARREN G. FINDLEY, ROBERT L. THORNDIKE, CHARLES L. LANGMUIR, TOBY OXTOBY, JOSEPH ZUBIN, ROBERT J. HAVIGHURST, GUY E. BUCKINGHAM, PERCIVAL M. SYMONDS.

DR. NORTH: I was particularly interested in the remarks concerning the relation of the pool to availability and supply. In Kentucky we probably have a notoriously low percentage of capable students going to college and also a very high rate of drop-outs. As a result of that condition, the Kentucky Association of Colleges and Secondary Schools made a survey of the reasons why students drop out from high school. The one thing that stands out in my mind is that the drop-outs in schools were not particularly marked by being in need of financial assistance or by low scholastic talent. More than 50 per cent of them did not consult with the high school administrative officers or teachers before they left, which, in our minds, points to the one crying need for better counseling and guidance service to these students who drop out.

CHAIRMAN DYER: Are there other comments?

MR. BRAGDON: I am connected with a study of the transition from school to college, which is being carried on by three schools and three colleges under a grant from the Ford Foundation. We are particularly interested in what happens to the able student. One thing that worries us about giving the talented special treatment or a special curriculum is that it may increase their sense of being different from other people. I wonder whether Dr. Havighurst is willing to generalize from his account of the able boy who wanted to be a grocery store clerk: does he mean that boys and girls from the lower income groups don't want to be considered different, and that they fear that higher education will make them exiles?

DR. GULLIKSEN: I would like to broaden one of Dr. Havighurst's comments. He emphasized the fact that in different social groups you may have different standards that will suppress the development of ability. It seems to me that similar effects may also occur on a national scale. As a concrete illustration, there has been a lot of discussion in quantita-

tively-minded circles (in the American Statistical Association, for example) about the development of good statisticians, the need for more people trained in mathematics and statistics. Helen Walker, in her presidential address a while ago, pointed out the difference in our attitude toward mathematics and reading—not by social castes but over the country as a whole.

You hear persons who are responsible for educational programs seriously urge that there are individuals who simply do not take to mathematics and, therefore, should be taught very little mathematics. It is asserted that such individuals are just not mathematically-minded and hence must not be asked to master a subject which is difficult for them. In our society persons are likely to say with some pride, "Oh, you have put that in figures and graphs so I can't understand it. I always pay attention to the commonsense of any situation."

If you look at the other side of the problem, what do you find? Do responsible educators say we should separate our population into those who are potentially literate and those who just cannot read? Students must study English each year all the way from first grade through college. There is a growing trend to provide special classes for those who cannot read adequately from grammar school right through college. You never hear anybody say with some pride, "Oh, you put that in words, and of course I can't be expected to understand it."

I think that difference in our total social attitude toward verbal and quantitative competence is reflected both by students and by educators. Such a social attitude might well be responsible for some undeveloped quantitative talent in our society.

DR. CRISSY: The analogy drawn between the specialization of platoon football and job specialization is an interesting one. When specialization has occurred, you intimate we get better performance in the more narrowly defined job. We are apt to get more scandals! I think the parallel in industry or in government service is if you simplify jobs enough you will get discontented "experts" because the job does not have the interest and variety to be challenging. Incumbents turn to things like having strikes or joining front organizations or anything else that will give them the interest which the job does not give. I am not saying that platoon football automatically is going to give more and more scandal, or that job simplification leads inevitably to discontent; I am simply making a case for job enlargement. I would rather have an all-round football player than a specialist who is discontented; similarly, in the Navy I would rather have an all-round sailor than a discontented specialist.

DR. DAILEY: Apparently there is more scandal these days in basketball than in football. The point is, if you are concerned with getting the high-

est level of expression of the talent, that is, in beating some other football team, I think it would be disastrous to try to operate with any system other than the two platoon system with that type of material.

DR. FINDLEY: I wanted to comment on one factor that tends to reduce the advantages of specialization. The greater the degree of specialization that one allows in individuals, the more coordination of their activity by others is required. If we are to get only a small increment in numbers, as Dr. Thorndike indicated by increased specialization, is that going to make up for the number of coordinators that we shall need with the skill in coordination that has to do with personal relations, and so forth? I have understood, and I think that we all understand, that administrators are rather highly regarded and highly paid because they have this ability. I wonder whether we are not multiplying another problem, when we reduce the general manpower problem by specialization, in that we have to have more coordinators with more skills of their type in order to make the whole system work.

DR. DAILEY: I would think that if, by increasing specialization we decrease the drain on our manpower pool for meeting these various demands—which may be for college trained people or otherwise—we actually make it possible to select at a higher level on each of the multiple dimensions of these generalized administrative personnel. Actually, the same thing that gives you the greater efficiency may also enable you to get higher level people in the broad talent jobs where they have to be leaders, very good verbalizers, have high mechanical facility and also have a great deal of physical vigor. We can select at a higher level on those multiple dimensions because there has been less of a squeeze on the manpower.

DR. DAILEY: I would like to point out one other factor in relation to Dr. Thorndike's remarks. These remarks concentrate attention on intellectual aptitude and mechanical aptitude which are more closely related than the others. But in the other types of physical conditions—perceptual motor aptitude, motivation and interest factors, and temperament factors—you get more independence, so the total gain would be more than you would get by considering intellectual and mechanical factors.

DR. THORNDIKE: I should hate to have on an important job somebody who was very high on motivation and very talented so far as motivation was concerned but who did not amount to much on any of these other dimensions. It seems to me that the other dimensions which you have there tend to be, in most instances, additional requirements rather than talents which can stand on their own feet. That is not entirely so, but I think it is, at least, in part the case.

MR. LANGMUIR: I have two questions for Mr. Dailey and Mr. Oxtoby. For Mr. Dailey: Is there any quanti-

tative estimate of the gain in our utilization that can be expected by work simplification? For Mr. Oxtoby: Is there any quantitative estimate of the utilizable and visible talent which is not being used because it is in the wrong place?

DR. DAILEY: Well, of course, in a military sense I suppose there is no quantitative estimate. We are just now preparing to do some research on job simplification. My own estimate is that it would be very considerable. The gain is an indirect one in that what you will do is to avoid having the man of some ability on a job where he is not being properly utilized rather than trying to get these isolated types of aptitudes that can stand by themselves. The primary concern in job simplification would be to avoid having on a job personnel with a high degree of talent in dimensions that are not very closely related to the job.

DR. OXTOBY: There is quantitative evidence that there is a lot of waste. We do not have any exact figures yet as to just how much there is. There certainly has been evidence that for any particular variable—intelligence, motivation, interest, and so on—there are many people that are high on those variables and who could ultimately expect to go into one of the high level fields but who drop out at a number of lower levels. There may be additional variables that cause them to drop out. I should say in answering that there is some quantitative data that there is a certain amount of waste but not any that say exactly how much there is.

DR. ZUBIN: My comment is directed to all three of the speakers. I am not directly concerned with the problem of selection of talent but with the problem of finding out why certain people of talent break down and do not do as well as they might. One concept we have found useful in studies of this problem is the concept of general structure of ability and personality. We have been very much disturbed by the fact that the general procedures of correlational techniques do not seem to apply when you consider these broken-down personalities. We have to depend—perhaps because we do not have enough cases or perhaps because we do not understand the problem too well—on the Gestalt or the pattern of ability and disability. We find some people who on the basis of their verbal, intellectual, and numerical ability—should be very gifted, but they do not perform on that level, and in fact, do not perform at all. On the other hand, we sometimes find people who are much lower in their general level of ability but who have certain assets which prevent them from becoming a drag on society.

I wonder whether in any search for talent and development of talent in general it is not essential to consider the totality of the Gestalt involved in some sort of pattern. Some people who are low in intelligence and low in numerical ability have a

good deal of friendliness and they can do a given job because they get cooperation from their fellows. This Gestalt structuring of ability and disability seems to be an important matter which has not been stressed enough by the speakers.

DR. HAVIGHURST: Mr. Zubin's comments could open up a pretty wide field for discussion. It does seem to me that, in the search for talent, it would be wise for us to go down, let us say, to about a 110 IQ. I am involved at the present time in a project where we are trying to identify all of the able—we call them gifted—youngsters in a community and identify them beginning at the age of ten. We are deliberately going to take as a first screening an IQ of about 110 and above on the theory that you need a substantially better than average general intelligence, but then we shall look for other factors and I am sure that the factor of a fairly well-balanced and stable personality is one of the main ones. I think that in picking out talent that is worth training, that is worth an investment by society, one should look for a kind of personality which shows a fair amount of stability and balance and toughness—I don't mean that one should require a kind of one hundred per cent extroverted American personality, that is not what I am talking about—otherwise we can get a high development of some particular ability which will not be socially useful.

MR. BUCKINGHAM: Is any work going on to study the idiosyncracies of the people to whom these people report, such as high school teachers, college professors, officers, or foremen?

MR. SYMONDS: I should like to follow up Dr. Zubin's remarks and also to raise a point with regard to something that Dr. Dailey said.

It seems to me that Dr. Dailey has assumed that these various factors or components are relatively constant in the individual, and although it is common practice to make that assumption with regard to intelligence, it is not safe to make it for motivation. Motivation may go up or down and change its pattern considerably. For instance, is there any guarantee that an individual would be as highly motivated when jobs are highly specialized as he would be when jobs are more complex?

Dr. Havighurst's case of the talented boy who resisted his parents' wish to send him to college even though they had the means interested me very much. In addition to the socio-economic factors, there were interpersonal relationship and emotional factors which undoubtedly prevent a number of gifted individuals from developing their talents and which, conversely, help some individuals to develop their talent to a very high level. I think that in this discussion we cannot neglect this factor of motivation and the emotional factors and interpersonal relationships that make it difficult for a boy to develop to his full capacity.

DR. DAILEY: I think that ties into my own point of view. I should like to point out that I am thinking of talents as types of dimensions and that ordinarily the job or profession will require some talent. I should like to tie that in with Dr. Havighurst's idea of talent as being the upper 2 per cent. An example might be the boy who is at the 97th percentile in IQ, so he is not talented if you consider that alone. However he might be in the upper 1 per cent on Dr. Bennett's test, or perhaps, have such a stable personality and such motivation that he is 'way up in the 1 per cent on that. He must be regarded as high level talent even if he is not in the upper 2 per cent in IQ. I also should like to point out that I said that I deferred to the educators and social scientists in the sense that they can probably do more than the personnel psychologists to increase the supply of talent when I pointed out as the fourth objective that of assuring each individual in the population of a full and adequate opportunity to develop and use whatever latent talent he has.

I think that there is a considerable amount of evidence that these environmental factors do a great deal even in influencing such things as scores on printed tests. When you get to such things as temperament and motivation, they probably do much more. I don't think that we are in basic disagreement on that.

CHAIRMAN DYER: I should like to announce that it has been brought to my attention that there are two recent manpower studies in which some of you may be interested. One of them is, "Employment, Education and Earnings of American Men of Science," Bulletin 1027, United States Department of Labor. The other is, "The Production of Doctorates in the Sciences, 1936 to 1948," a published report of the project sponsored by the Manpower Branch, Human Resources Division, Office of Naval Research, and it is published by The American Council on Education—an imposing looking volume.

These two studies can be had by sending your request to Ralph M. Hogan, Manpower Research Branch, Human Resources Division, Office of Naval Personnel, Navy Department, Washington, D. C.

ADDRESS

PAUL L. DRESSEL

PROBLEMS OF EVALUATION IN GENERAL EDUCATION

CONSIDERED FROM the viewpoint of an evaluator, the general education movement evinces two different and frequently conflicting aspects. The first of these and the one as yet most characteristic of ventures in the development of general education programs is the organization of courses sampling widely the subject-matter of a number of related fields. Such titles as Social Science, Biological Science, Humanities attest to this trend and are indicative all too frequently of courses placing emphasis on the inculcation of the essential elements—usually facts—of the cultural heritage. Evaluation in such a context poses no particular problem, for the task is essentially that of testing factual knowledge. The tests utilized in most such courses demonstrate this characteristic. The one difficulty faced is that of constructing a test which is generally applicable to schools concerned with general education; for the wide variety in selection of materials results in it becoming a practical impossibility to write a test applicable and satisfactory to all. It is perhaps not unfair to say that the Tests of General Education developed originally by the Graduate Record Examination Office and the General Culture Test of the Cooperative Test Service, both now available through the Educational Testing Service, were developed on the basis of this first aspect of general education and have faced the difficulty just mentioned.

The second aspect of general education is a more complex one. It involves on the whole a concern both with the *use* and the *usability* of the material presented in a course. It involves a concern with use because the material selected—the content, if you please—may be selected, for example, because it is conceived of as conducive to illustrating and developing various intellectual skills sometimes subsumed under the title critical thinking or because it may modify the attitudes, beliefs, or philosophy of the student. This concept of use also brings into consideration the method of instruction utilized. Many teachers of general education have come to suspect that the development of critical thinking or changes in attitudes result as much or more from the type of classroom activity as from

the material studied. This second aspect involves also a concern about usability because the content may be selected in terms of what an educated individual can and should use in fulfilling his vocational, social, and community responsibilities or in enriching his own personality. It is apparent that this second aspect of general education faces the evaluator with many difficult and as yet unsolved problems. The immediate difficulty of deciding how to define operationally such popular phrases as critical thinking, understanding of the scientific point of view, effective communication, desirable attitudes, philosophy of life, good citizenship, personal adjustment, and the like, is by this time familiar ground to most measurement experts. The further problem exists that general education in this second aspect is not a series of courses, not the achievement of a series of independent objectives but rather the characteristic of an integrated personality functioning adequately or better in all aspects of living. The evaluator who proposes to solve this evaluation problem must, perforce, ask the Almighty to find him space on the Judgment Seat. Even if successful in attaining such sponsorship, the evaluator may still expect some skepticism in his dealings with those professors who might well be inclined to regard Omniscience as a quality obtainable only through acquiring a Ph.D. in their particular field.

In all honesty, it must be admitted that teachers have more reason to distrust evaluators. The mere presence of an evaluator involves an implication that teachers have left something undone. Even teachers who are definitely concerned with improvement in teaching and evaluation are likely to be hesitant or suspicious of the evaluator because:

1. Objectives have all too frequently been thrust upon the teachers by administrators or evaluators. In such cases objectives mean little to the teacher.
2. In dealing with separate objectives evaluation seems to segment the individual. Such analysis may be followed by synthesis but this procedure does not satisfy those who dislike and distrust analytical procedures.
3. To some teachers—particularly those in humanities—certain objectives have an esoteric, mystical quality the essence of which must be lost in any attempt to get it on paper. To such individuals the mere existence of a test, inventory, or any formal evaluative technique is prima facie evidence that it tests nothing of real significance.
4. A common tendency is an unwillingness to compromise with anything less than perfection. It matters not that teachers' judgments or tests are inadequate; the proposals of the

evaluator have flaws and are not to be accepted until these are eliminated.

5. Evaluation is regarded in the same light as testing for the purpose of determining a grade. At this point everything worthwhile has been accomplished and extensive testing is a waste of time. Evaluation is not seen as a basis for modification and improvement of classroom practice.

The preceding list of instructor reactions could be expanded almost indefinitely by listing other points or by enlarging on those already presented. There is no need for this, however, for these will document the contention of most evaluators that any evaluation activity which is to be truly productive must be so conducted as to be intelligible and acceptable to teachers. The simplest way to do this appears to be that of making evaluation a cooperative activity with teachers.

It is not to be inferred from this stand that evaluation is to stop with development of evaluation techniques and collection of data on student performance. In connection with such an objective as critical thinking, a great deal of fundamental research remains to be done. It is necessary to investigate such questions as

1. Is critical thinking a teachable ability or is it largely the spontaneous interaction of intelligence with knowledge?
2. Assuming that critical thinking can be developed, is it largely specific to the context in which it is learned or is it a generalized skill?
3. What teaching practices are most conducive to the development of critical thinking?
4. What is the relation between attitudes and beliefs (the affective aspects of behavior) and critical thinking?

There can be little doubt that the future of general education depends to a great extent on such research. Much of this research can be done more efficiently by research workers unencumbered with the necessity of operating in the slow moving cooperative situation. However, the fact that many generally accepted principles of learning are completely ignored in teaching—even by psychologists who teach the principles—is evidence that research in which teachers do not participate affects them but little.

The state of mind conveyed by the preceding remarks has grown out of my experiences in carrying on evaluation activities in the Basic College at Michigan State College and has been reinforced by recent experiences in the Cooperative Study of Evaluation in General Education. Indeed, the remarks may be considered as a rationale for both the existence of the Cooperative Study of Evaluation in General Education and the mode of operation utilized

in the Study. This Study arose out of a recognition by such leaders as Dr. Earl McGrath, Dr. T. R. McConnell, Dr. Ralph Tyler, and the late Dr. George Zook of a need for investigating the nature of certain objectives of general education and the extent of student development with regard to them.

In the fall of 1949, Dr. Zook, then president of the American Council on Education, invited me to survey the interests of a group of colleges and universities in participating cooperatively in an evaluation study. As a result of the demonstrated interest, coupled with a willingness to make some contribution to such a project, the Committee on Measurement and Evaluation of the American Council on Education invited 19 colleges and universities* to join in the venture. Exploratory meetings and college visits were carried on during the period from December 1949 through June 1950. This period was devoted to ascertaining a set of common objectives and related projects in which all the colleges would engage. It is important to note that the cooperative effort was directed toward the pooling of resources in dealing with the more significant common problems faced in general education programs, rather than to dealing with unique institutional concerns.

Active work in the project dates in reality from the August 1950 workshop in which six inter-college committees met for two weeks in East Lansing. These committees, dealing with objectives in science, social studies, communications, humanities, critical thinking, and attitudes, devoted most of the two weeks to developing operational definitions of objectives and to sketching the specifications for the preparation of evaluation instruments. As an illustration, the definition of critical thinking reached by the Social Science Committee involved the ability to

1. Identify central issues in a problem.
2. Recognize underlying assumptions in a statement.
3. Evaluate evidence or authority.
4. Draw warranted conclusions.

In addition to this definition decisions were reached then or later as to the relative weighting of each of these in the proposed tests and as to the content areas which would be sampled in developing test items. The content or problem areas included (1) the culture concept, (2) economic affairs, (3) political affairs, (4) so-

* Antioch College
Boston University
Colgate University
Colorado State College of Education
Drake University
Florida State University
University of Florida
Harvard University
Kansas State College
Kansas State Teachers College
Michigan State College
University of Minnesota
Muskingum College
Oklahoma A and M College
Pennsylvania College for Women
Stephens College
Syracuse University
Western Washington College of Education
Wright Junior College

cial affairs and were obtained by a survey of the materials included in all the courses represented in the committee. Some sample items and questions illustrating each phase of critical thinking were prepared to guide the committee and faculty in the preparation of test materials.

Certain rather significant conclusions emerged from this initial workshop experience:

1. General education courses in a given area—social studies, for example—were found to have a far larger common body of content than had been suspected in advance.
2. It is possible for teachers to agree upon and to put in concrete and operational terms the meaning of such vague concepts as critical thinking. Furthermore, the definitions arrived at by teachers of social studies agreed in essentials with definitions independently developed by teachers of science, humanities and by psychologists and philosophers.
3. The various objectives of general education are so interrelated that the complete analysis and evaluation of any one of them inevitably involves all the rest. Thus, for example, attitudes and knowledge are involved in critical thinking. Likewise, good citizenship implies the ability to think critically.
4. Evaluation is effective only as it makes possible both the measurement of changes in the student and the improvement of the instructional process aimed at producing these changes. Some members of the committees came to evince as much interest in the development of exercises and techniques for improvement of class instruction as in the development of evaluative instruments.
5. Fundamental research is needed in regard to the learning process. Such questions as: How do the critical thought processes of the student develop? Can critical thinking be taught? How do attitudes impede or promote critical thinking? Is critical thinking specific to the courses or area in which it develops or does it carry over into other areas?

Most of these conclusions or questions were no surprise to the Study staff or to the Committee on Measurement and Evaluation, but it was significant that teachers should arrive at them through their own deliberations.

During the academic year 1950-51 efforts were directed to the development of a variety of tentative instruments and to extensive tryout of them. At a workshop in June 1951, the committees with the assistance of competent evaluators in each area and with data from the tryouts, produced a number of instruments for use during the year 1951-52:

COMMITTEE ON ATTITUDES, VALUES, AND PERSONAL ADJUSTMENT

Inventory of Beliefs Form I: Designed to test attitudes toward a number of commonly held beliefs and through them to obtain evidence about underlying personality structure. This inventory was built around the use of common misconceptions illustrated by:

1. Lowering tariffs to admit more foreign goods into this country lowers our standard of living.
2. Any man can find a job if he really wants to work.

Its rationale was based to a considerable extent on the F scale of Sanford, described in the book called *The Authoritarian Personality.*

Problems in Human Relations Form I: Designed to obtain, by means of students' responses to conflict situations, evidence about their attitudes toward the sanctity of human personality. Situations culled from student writing and interviews and from faculty observations were used as the basis for statements about solutions, appropriate courses of action, or analysis of the basic issues. From these statements the student was instructed to choose that which most appealed to him.

COMMITTEE ON COMMUNICATIONS

Critical Analysis Test: Designed to obtain by objective means evidence about students' reading ability and knowledge of writing techniques. In addition to this test, the committee has developed a scale for judging communication effectiveness based upon a critical incident approach and emphasizing the common characteristics underlying all forms of communication. They are also developing a set of themes carefully rated by a committee for the purpose of encouraging greater uniformity in theme reading.

COMMITTEE ON HUMANITIES

Humanities Attitudes Inventory: Designed to elicit evidence of students' feelings, either for or against, a number of aspects of the humanities such as kinds of works and those who are employed in the humanities. This inventory is simple in format, requiring only a strongly agree, agree, disagree or strongly disagree response to such statements as

1. The arts are not important to society.
2. The reading of plays is boring.

Humanities Participation Inventory: Designed to obtain from students an indication of their interest, experience, and participation in a number of humanities-centered activities. This is also a very simple instrument which requests the student to check the appropriate statements in a sequence illustrated by the following:

You read poetry

1. (1) never
   (2) occasionally
   (3) frequently
2. (1) with little or no enjoyment
   (2) with a fair amount of enjoyment
   (3) with great pleasure

3. (1) for the moral, the sentiment, or the rhyme
   (2) recognizing such elements as similes, metaphors, rhyme schemes, dominant ideas, etc.
   (3) making a careful examination of the formal elements, structure, ideas and their inter-relationships.

The first group involves simply contact with the activity (frequency), the second group involves the enjoyment derived and the third the seriousness or thoughtfulness of the approach to the activity.

A Guide for Critical Judgment in the Humanities: Three separate forms, each designed to obtain evidence as to students' ability to make an analysis of a work in the humanities. Forms for Painting, Music, and Literature have been developed. This test involves a structure which calls the attention of the student to certain characteristics of a work of art such as:

1. Their subjective reaction to the work.
2. The function and context of the work.
3. The nature, use and importance of the medium in the work.
4. Formal elements of the work.
5. The organization and style of the work.

It requires an essay response to each plus a final evaluative essay response involving all these elements.

COMMITTEE ON CRITICAL THINKING

A Test of Critical Thinking Forms A and B: Designed to be comparable forms of a test to measure critical thinking ability without reference to any particular body of subject matter. The tests involve a series of problems accompanied by statements requiring various aspects of critical thinking.

COMMITTEE ON SCIENCE

A Test of Science Reasoning and Understanding Forms A & B: Designed to measure students' ability to read and interpret scientific materials written in popular style, these tests utilize carefully selected material and seek to find out the extent to which students can analyze, explain, or criticize statements, hypotheses, or conclusions presented.

COMMITTEE ON SOCIAL SCIENCE

A Test of Critical Thinking in the Social Sciences: Designed to measure students' ability to think critically in the area of the social sciences without requiring a high level of specific knowledge of social science content. *A Vocabulary Test* for use with *A Test of Critical Thinking in the Social Sciences:* Designed to aid in finding out the part knowledge of content plays in the critical thinking process. This test involves social science vocabulary actually used in the critical thinking test in social science with the hope that poor performance in critical thinking which is largely due to inadequate vocabulary can be distinguished from sheer inability to think.

Each committee, with the assistance of a Research Committee composed of research directors and evaluators, developed a series of studies designed to throw further light on the validity and reliability of the instruments as well as to give some data on student development and to throw some light on a number of fundamental questions or hypotheses such as those raised earlier. (See point No. 5). All of the tests and testing programs are being given extensively in the cooperating colleges this fall. We shall remark on the purposes and use of the results shortly.

Assuming that continued financial support is forthcoming, it is proposed that the Evaluation Study continue until January 1, 1954. The work of the study would embrace the following activities:

1. The accumulation of data on the validity, the reliability, and the usefulness of the various instruments and probable revision of some of them as a result of this information.
2. The accumulation of data on the initial and final status of students at the beginning of their college career and at the end of periods of one year and two years respectively.
3. The tentative investigation of a number of hypotheses about the nature of the general education objectives measured by the instruments, including such items as their inter-relationship and the relative effect of various educational experiences in promoting student development with regard to the objectives.
4. The preparation for publication of a final report of accomplishment, of research findings, and of suggestions for further activity based on the experience of the Study.

It is fully recognized that such investigations as the collection of data on student progress and other research are dependent on the validity and the reliability of the instruments used. It is also true, however, that the findings of such investigations are relevant to the determination of validity. The urgency of the situation suggests that the simultaneous attack on all of these fronts will yield valuable information which, even though somewhat tentative in nature, will bring the project to a close at a point where the individual institutions may carry on further research. Furthermore, another and more specifically research project may be built on the conclusions reached in this one.

The need for good evaluation instruments in general education is obvious and the need for developmental studies rather than status studies is equally obvious. It is these that are embraced in activities 1 and 2 above. However, the research implied in the third activity is, if anything, even more fundamental and deserves further clarification. Six significant pro-

jects are envisioned in this research area. They are:

1. A study of growth and achievement in social sciences, as measured by the test of critical thinking in Social Science, in relation to performance on the Test of Critical Thinking and on the Tests of Beliefs and Attitudes.
2. A study of the relationships between certain kinds of attitudes and certain traits or abilities involved in the field of the Humanities.
3. A study of the relationships among the three Humanities instruments and the Tests of Critical Thinking.
4. A study of scores on the Reading and Writing Test related to scores on the tests of Critical Thinking and the Attitudes Tests.
5. A study of growth in scientific reasoning and of factors related to such growth.
6. A study of the nature of and the factors related to growth in the ability to do critical thinking.

A brief discussion of the first of these will make clearer the intent of that particular project and likewise will clarify the type of research involved in all six. The Critical Thinking Test in Social Science involves critical thinking in the context of college level social science materials. The Critical Thinking Test involves the use of only such facts and principles as are believed to be in the possession of most entering freshmen. The Inventory of Beliefs attempts to assess the underlying biases and the resulting rigidity of individuals possessing them. The Problems in Human Relations Inventory is designed to determine whether students accept democratic or undemocratic procedures and solutions in situations and problems chosen so as to be as realistic as possible. In addition to these tests, data on intelligence and, in many cases, on the extent of the student's social science vocabulary will be available. By careful selection of groups tested and by the proper relating of these data, the following questions can be investigated:

1. What is the relation of intelligence and of vocabulary knowledge to improvement in critical thinking?
2. To what extent is improvement in critical thinking dependent upon the possession or the acquiring of certain beliefs and attitudes?
3. How is improvement in social science critical thinking related to improvement in the less specifically content-oriented critical thinking and to what extent is this conditioned by the pattern of courses carried by an individual?
4. An extensive biographical data sheet available on each student will make possible—if time and money permit—the study of

such factors as cultural and economic background in relation to such changes as mentioned in 1, 2, and 3.

It is apparent that the possibilities are almost limitless and the task is largely that of selecting the more significant ones for immediate study. In this we have sought and continue to receive advice from a number of research directors who have served as consultants to the project.

At the conclusion of the Study, the findings to be summarized in one or two published volumes would include:

1. A description of the origin, the mode of operation of the project, and the place of evaluation in general education.
2. The rationale for the selection of objectives studied and the definitions of these objectives which were arrived at.
3. The rationale of the various instruments developed and the presentation of evidence on their validity and reliability.
4. Considerations of certain fundamental questions about the interrelationship of objectives and the extent of student development in regard to them. This would involve the research aspects just discussed.
5. Indication of some of the issues not touched upon or of those needing further study. In short, a summary of progress made and a series of suggestions as to next steps.

It would also be hoped that some if not all of the instruments might have such value that the Educational Testing Service would be willing to publish them or to incorporate them in its proposed *Tests of Important Educational Objectives*, thereby making them generally available. This availability, coupled with the evidence and suggestions made in the Study reports, should effectively stimulate similar evaluation and research practices in general education. Such stimulus is vitally needed if general education is to continue its forward progress and give more than verbal support to the aims which characterize it.

DISCUSSION

PARTICIPANTS:

HENRY S. DYER, DOUGLAS E. SCATES, PAUL L. DRESSEL, CHARLES R. LANGMUIR, HELEN JENNINGS, HERBERT S. CONRAD.

CHAIRMAN DYER: We have a few minutes for discussion. Are there any questions or comments or criticisms?

DR. SCATES: To develop critical thinking in the field of humanities and social studies, I wonder whether any attention was given not only to the question, "Is this true?" but also to, "How much do I want it?" and, "How desirable is it to me as an individual, for the group of people that I go with and, of course, for society at large?" In other words, in the social studies and in the humanities we are, to a large extent, dealing with batteries in our teaching and I wonder whether critical thinking, as we seek to measure it, has been spanned to include these elements.

DR. DRESSEL: I am not absolutely sure that I get the point of your question, but let me comment on it as I understand it. I think that the committees that were concerned with critical thinking did recognize that there was some question as to whether students were really motivated to develop this particular kind of skill. They felt that in the test they were developing they ought to deal with critical thinking pretty much as an intellectual consideration, leaving to the people who were developing the tests of attitude, to which I referred, to work into that pattern something which would give some impression as to whether the student had prejudices against this sort of thing or whether he had biases that would interfere with critical thinking. It was discussed extensively, but there was a feeling that more tangible results would come out of different instruments to get at this than to try to work them into the same instrument.

MR. LANGMUIR: In earlier cooperative studies there have usually been news letters, interim reports, or bulletins which have been available, to non-participants. Is there any arrangement in this study whereby non-participants may keep in touch with the progress, obtain copies of the test materials, see the preliminary data, and so on?

DR. DRESSEL: I am afraid that I must answer by saying that no par-

ticular arrangements have been made for that sort of thing. The way in which we were operating and dealing with objectives led all of the committees to feel that they wanted to keep this a kind of closed corporation. They did not like the idea of these materials getting out until they were at a stage in their thinking where they thought they were reasonably respectable. I think we probably are at that stage now. In fact, a few people outside of the study have some copies. The gentlemen sitting next to you, Mr. Remmers, obtained some copies of it this fall and a few other people have written in. Where they are people that we know have a background in the field, and where we can give them a little understanding of the way in which we are working, we are willing to do so. There is no general program set up for passing around such information outside of the study as yet.

MISS HELEN JENNINGS: I wonder whether there has been included in the study any similar testing of the people doing the teaching of the students tested, that would give a background of at least the group membership so one could situationally infer what their values are, if they belong to teachers unions and things of that sort, and similarly, for the parents of the tested students. Has father or mother worked for any cause, and if so, what cause?

DR. DRESSEL: This is a cooperative enterprise so people do, to some extent, what they wish to do and, in general, they do not wish to be taking these tests themselves. However, to some extent, the thing to which you refer has been present. The way in which we function has called for a number of intercollege committees with one person from the institution sitting on the committee and, in most institutions, a larger number of people at home cooperating through the agency of this particular individual. Most of the tests which have been developed were, in their tryout stages, taken by almost all of the home staff in these various institutions, and their reactions and experiences in taking them were utilized in connection with revision but not in terms of studying the faculty members themselves. We had a proposition made by Dean Morris of Minnesota last June—he made it twice, in fact, at two of our general meetings—that everybody in the six intercollege committees be required to take all of the tests developed, but this was voted down both times by a good solid majority.

MR. CONRAD: I was a little disturbed at what seems to be a somewhat dogmatic assumption, in both the morning session and this session, with regard to the need or advantage of a well integrated or well balanced personality. It seems to me that what is wanted is training that would lead to some useful accomplishment. As I see it, Henry Ford inventing a cheap automobile did not require what is meant by a well balanced integrated personality. Henry Ford, as

the administrator of a great enterprise, probably would have been a better man, socially had he had a better balanced personality. In other words, for the work of the administrator, a well integrated personality may be necessary for good social accomplishment in that work. For more specialized work, we may not require a well integrated personality. We may, perhaps, require a concentration of motivation bordering, sometimes, on the fanatical. I might say that sometimes the Wright brothers had that kind of concentrated attention for a limited field of activity. In any event, I do not think that we ought to set up something as a goal unless it is clearly demonstrated that it is useful to the work for which it is assumed to be necessary.

LUNCHEON ADDRESS

MAJOR GENERAL LEWIS B. HERSHEY

CHAIRMAN DYER: Ladies and Gentlemen, I don't think it is conceivable that anyone here does not know the president of Educational Testing Service but, just out of chance that they do not, may I introduce to you my old friend and the president of Educational Testing Service, Mr. Henry Chauncey.

MR. CHAUNCEY: Some years ago, in the era of the Model A Ford, two friends of mine took a trip through Nova Scotia, in a blue Model A Ford roadster. It so happened that a rather unusual event occurred in Nova Scotia—it would not have been unusual in the United States, but it was quite unusual in Canada. There was a bank robbery. The thieves made their getaway in a blue Model A Ford roadster. You can well imagine the effect on my friends' tour. They were stopped some eight or nine times, hauled into the police station, and made, at great length, to explain their presence and give alibis as to where they had been at the time of the robbery. Eventually they would be cleared and allowed to proceed.

One evening, just at dusk, they were driving along the road beside a field. It was in that twilight time when things are somewhat deceptive. There was a field of grain on one side of the road. A slight wind was blowing and the grain was moving with the wind. The fellow who was driving looked over and said, "My, isn't that a beautiful lake?" His friend turned and said, "That is no lake; that is a field of grain." He said, "Why it is a lake; you can see the waves." His friend said, "That is a field of grain; you are crazy." The driver said, "Why it is a lake. I am positive it is a lake; you can see the light shining on the water." "Well," his companion said, "all right, I'll make a bet with you. If it is a lake, I'll dive into it, and if it is a field of grain, you dive into it."

So they stopped, got out and examined it, and found it was a field of grain. The driver walked out and did a magnificent swan dive into the field of grain.

It just happened that, at that point, some of the Nova Scotia police who had been following them came up. They stopped and came out into the field to inquire what was going on. As my friend later said

to me, "If the truth ever sounded foolish—I dove into a field because it wasn't a lake," whereupon they went off to jail again!

The truth does not always—or even often—sound foolish, but it does frequently sound surprising. I think if it had been said a year ago that the man who would make the greatest contribution to the advancement of testing in 1951 would be General Hershey, the Director of Selective Service, that would definitely have been surprising. Yet today, I am sure that all of you would agree with me that, due to the draft deferment program which included a test as one of the bases for deferment, it was an important contribution to the advancement of testing. For two reasons: one, because it is the first occasion when tests have been used as an instrument of an important national manpower policy and secondly, because it aroused a great deal of interest in testing on the part of many, many people who had not had a great deal of interest in testing before that time. Millions of new people were going to take the test, and other millions of people were concerned about the general problem of the use of the tests. It aroused a great deal of interest and discussion.

I think it is most fitting that today we shall have a special luncheon—such as we have not had before at these conferences—and have as our guest General Hershey. I am actually more interested in having this luncheon and having General Hershey here for another reason and that is simply that I should like you to have an opportunity to get to know General Hershey. I know that every one of you is thoroughly familiar with him as a public person whose name appears in the press constantly, whose sayings are quoted in the press, and who appears in newsreels, television, and radio programs. I think it is much more important to know him as a man, and I think you will be able to do so today.

General Hershey is, to my mind, a most unusual public servant. He handles a very difficult and extremely important job, one that by its very nature should be perhaps the most unpopular with the public. It is the kind of job in which you might expect a person to stay a year or two and then, under pressure, withdraw. General Hershey has been Director of Selective Service for something more than ten years, I believe, and during that time he has gained the complete confidence and respect of everybody from the President of the United States down to the lowliest draftee.

What manner of man is it who can accomplish that very unusual feat? In the first place, he has amazingly good judgment compounded from high intellectual ability—he would score very high on any of our tests—and common sense. Common sense is a quality which we have not yet adequately measured, but one which I think we all regard as important and one which I am sure we would rate extremely high. He

is an idealist and he is farsighted. He sees important objectives in the distance and sets his course towards them. At the same time, he is a strong believer in the democratic economic process at all levels, from the lowest to the highest, and he realizes that goals cannot be accomplished immediately. He has, therefore, patience and an unusual philosophic attitude—but not the philosophic attitude that is passive and just throws up its hands if it is not accepted. He has, rather, the philosophic attitude that also embodies doing everything that he can to see that important objectives are achieved. And he is an extremely hard worker.

One day last spring when I was going down to Washington on the train, I ran into General Hershey just as we were getting off and he kindly offered to give me a lift to my hotel. I noticed he was carrying a bulging briefcase, and I noticed that his chauffeur brought another bulging briefcase for him to take home that evening and go over before the next day's work.

He has in addition, I think, a real and genuine love of people and what Kipling would have called "the common touch." Supporting that, he has a marvelous sense of humor. To attend a meeting with him, even when serious and important matters are being discussed, is always a pleasure. Dispersed here and there throughout the meeting are touches of humor that make it extremely enjoyable as well as importantly worthwhile. I remember that on one such occasion the subject of speeches came up. He turned to me and said, "You are probably like me; you have just one speech. Sometimes I begin at the beginning and go to the end. Sometimes I begin at the end and go to the beginning, and sometimes I begin in the middle and go both ways. That confuses them, but they think it is profound."

I don't know what General Hershey is going to talk about today and I don't know in what direction he is going to go through his subject, but I am sure that you will find that it is profound and very interesting. It is my great pleasure to introduce to you Major General Lewis B. Hershey.

MAJOR GENERAL LEWIS B. HERSHEY: Introducer, Chairman, Ladies and Gentlemen: I have a sense of humor, or I would probably believe more of that. I have a great deal of satisfaction in finding this is a luncheon, rather than a funeral, and that I am still alive. I did not expect to hear anything like that while I was alive and, of course, obviously I did not expect to hear that after I had quit living.

I have been a little nervous about coming up here. In the first place, I haven't any speech left—he told my story. I guess the copyright was out on it, but, just the same, it was my story. The next thing is that I am facing individuals who are probably at the present time engaged in making

some sort of a recording. I don't know that you can go any place and be declassified long enough not to be professional by attempting to test at all times, especially when you get together at one of these meetings.

I have been awed from afar by the things that people in the testing world have been doing for the last 50 or more years. I know most of you men have been at it about that long, and I realize the ladies in the audience have only recently started. It is not because they are not interested in testing; it is just because they have not lived as long as these loose hairs that are turning grey. So I have been awed by the question of testing. However, I thought it a sort of prosaic thing, and I was not prepared today to be caught by the poetry of Mr. Chauncey, even though he had to make Canada the locale and introduce some of the things that we see over the television —bank robbing for example—but, after all, the wheat waving (not unlike water) was nothing like poetry.

I have viewed your work through the eyes of a very rank amateur over a very long period of time. I am probably one of those persons who suffer with either double or triple frustration. When I was young I probably was not educated to be either a lawyer or to get into this testing, or to be a soldier, and I have never gotten to be any of them; therefore I am suffering with double or triple frustration. I have had, and do have, a very deep feeling that the area in which you are operating is probably the greatest frontier that we have. I am not thinking particularly of the testing; I am thinking of the ignorance that we meet on this frontier or the ignorance that you are trying—I hope incessantly—to penetrate, because I think the frontier of our lack of knowledge, especially about ourselves, is the greatest frontier we have, and I am wondering whether there is as much lag between the time of finding out what we do not know, and the arrival at knowing, as there has been between the first synthetic striking of fire and the first fission of the atom. I don't know what threshold we stand on, but I have had a great deal of interest. I probably hoped at one time that someone would know what kind of persons we need as officers in the armed forces. Perhaps I am not so much interested in that as I was ten or fifteen years ago. The problem has been one of determining how to grade what we had, when we had perhaps very vague ideas of what we wanted. Sometimes I saw that the more we codified what we knew, the less we understood and, therefore, the more we put down on our efficiency reports perhaps the farther from actuality we got.

I am not criticizing anyone. It is merely the fumbling that goes on in trying—either with words or even with pictures—to portray what goes on inside of the human being or somebody that looks like one. Therefore, I have had a very deep respect

when I have looked at you from a distance. I cannot say that I have not more or less retained my respect as some of you have come a little closer, but that is not the purpose of the things that I should say today. I should like to come up here and straighten you out on how to run your business. That is not because I know anything about your business; I don't. Anyway, the people who come down day by day to help me run my business—well, many times I think they know about as much about my business as I do about yours. Also, I am 'way behind on giving advice to other people. I am 'way behind on carrying out advice I have gotten from other people. In fact, it is so stacked up I don't know whether I ever shall catch up.

Approaching any question we have in this country, because I am not a scientist—I studied liberal arts in my day (I don't know what they call it now) they didn't always call it that, but I mean when you spoke of it as you should, you called it liberal arts. So having come out of the realm of liberal arts, I am under some embarrassment when I get around to scientists. But I like to feel, in as much as I don't know any too much about any one thing, that the trouble with the world perhaps is that we have too much compartmentalization of our knowledge. I am confronted in the manpower field first with the manpower problem, which in itself is merely a part of many other things. But within the manpower problem itself we have the problem of the skilled man; we have the problem of the engineer; we have the problem of the scientist of one kind or another, but there is only one manpower problem and that is the whole, and to solve any of it gets you immediately involved in all of it.

There is about only one thing wrong with manpower, we are too short in both quantity and quality; otherwise the manpower problem is relatively simple. If we are going to use manpower narrowly, it is pretty nice to think how rare we are and how short in supply we are. Just think of a world getting to a place where it is in short supply of such as we are.

You like figures; I don't, but I'll mouth a few of them. How accurate they are is always a question, but how accurate many of your figures are is many times a question. There are over 150 million people in the United States. As I was coming up here from the train it occurred to me there were more, but perhaps the streets I was on were a little heavier populated today. I don't want to confuse you too much, so I am going to retreat immediately into about 12 million 4. Those are the individuals who are between 18 and 26 or have been during the last few years who are known colloquially as registrants. That is not too dangerous if you do not get too highly specific in the registration field. With three sentences I can complicate our problem considerably. Three million four

hundred thousand are over 26 and are no longer liable; that leaves 9 million. Two million have already gone in; I suppose we can get them in again if they got out sometime but until they do we can't do much about them, so that is 7 million. Two million of them are veterans under 26 years of age and are not at the moment liable. So with three sentences I am down to 5 million out of 12½ million. That is about the way it goes. It is like money on payday. Of this 5 million, 1 million are 18 years old, and while half of them become liable, they are only partially liable because you cannot take anybody on under 19 unless you can establish by scientific means of testing that you have no one liable that is available over 19. Therefore, while this million has half a million of productive units gross in it, there is some restriction on their use.

I am down to 4 million now. There are a million, in fact a million and one or two hundred thousand, who up until this summer thought they had dependents, and dependents which prevented them from serving in the armed forces, but when Congress passed the 1951 Act in June it took a slightly different approach. Congress required that they have more than one dependent. One dependent did not count anymore. Well, that is correcting itself rather rapidly. But, even so, I think we might find enough to gross around a hundred thousand or so—we did think two hundred thousand—but that was last summer and time changes all things and we may get a hundred thousand gross—of which perhaps thirty or forty thousand can pass the physical examination, if they get so far, if they cannot prove that they are farmers or tool makers or some other reason that may defer them. I am down to 3 million now.

Then we get into a miscellaneous group of 6 or 7 hundred thousand, such as a couple hundred thousand or 250 thousand all the time, whom you are going to have unclassified because they are just moving into the 19-year-old group. You are going to have a couple hundred thousand, and probably more by spring, who are deferred because they are students, or at least they have evidence to indicate that they are. Then there are a few other groups, such as ministers and conscientious objectors and aliens and what not. Eventually we arrive down somewhere near 2 million, which can be divided into two parts, although there is lately more than a million in one of the groups, and those are the ones that we have in 1-A, and of course are the fodder we are feeding just at the moment—if we want to use that term—and the other million plus are the ones we have had up but did not pass. Doctors turned down some of them because they had feet off, some of them were turned down by psychiatrists because they had parts of their heads missing or super activity in some of the parts that were there, and then there is the goodly share of them that

are either too stupid to pass the Armed Forces Qualification test—or smart enough not to. We lose them in either case.

So we have around a million we are working on at the moment. We are supposed to have approximately two hundred thousand who have been accepted. That doesn't mean they can be inducted because the doctors change their minds—after all to change their minds is human—but we lose 8 per cent in September just on the differences of opinion between those who conducted the pre-induction examination and those who conducted the final examination. So we shall lose some of the two hundred thousand already accepted, and we shall be lucky if we don't lose somewhat more than half of the eight hundred thousand who have not been examined.

There we are. I think, or did think up to a few weeks ago, that this physical situation in manpower was going to begin about Christmas a year from now. When we started business in July a year ago we had two jobs; we had to build an armed force and we had to maintain it. We are pretty well on towards the building. We are just starting to maintain it. It is like marriage. As has always been said about a marriage, it isn't the cost of acquisition, it is the cost of upkeep. So it is with maintaining the armed forces. Many times it is easier to build because there is more enthusiasm and all that goes with enthusiasm. Not only that, you can use two kinds of manpower; you can use the income and you can use the surplus if you start with one, and after you have been in the maintenance business a while you can't use the surplus for a fairly good reason. It is used. It is like the fellow who said to the girl, "May I have the last dance?" She said, "You have had it." That is the way with our surplus; it is not quite all gone, but it is going. You people could go at this in many more ways than I can because I have no scientific approach, but we had one million six hundred thousand in I-A July 1, 1950.

We are down now to right around a million and we have not only used up six hundred thousand of surplus, but we have used up the numbers that would be coming in during the last twelve months. Now it seems to me that there is no question that we are going to run out of surplus sometime. The time will depend on whether the interim developments tend to give us more men or give us less. It is a strange thing—and I think it is a wonderful thing to show we are always optimistic—we always think we are going to have more money at the end of the month than we ever do have. The same thing with manpower. We always think that the pool of manpower is going to be greater at the end of any period than it is. There are many reasons for it.

Our armed forces, as we built them a year ago, by necessity had some weak points. That is because of individuals. A great many individu-

als were called into service from the voluntary or involuntary reserves who had hardly made plans, didn't know where they belonged. They came in for somewhat of a surprise. They found they belonged to an organization. Congress took interest in the fact that they had been veterans and placed limitation on their length of service, which, by the very nature of things, would tend to reduce their numbers in twenty-four months, and after you reduce part of the individual's twenty-four months, seventeen or something like that, and he is discharged at twelve or a little after, then you have elements of instability in the numbers that you have acquired and you don't hold your numerical strength so well if numbers of the men are going out ahead of time, which they always do. And there are always men developing hardship and all sorts of things. You always lose more than you think you are going to lose. That is one reason why, when I mathematically compute, I always substract two or three or four months calculations. Something will happen to upset the calculations at least that much. Our rejection rate has been rising rather rapidly. I know it won't go up to the ultimate, but the ultimate would be one hundred per cent.

I realize we probably won't reach it, but when it gets around to fifty per cent I am not so sure that that is not where we are heading. I do not believe that I am being overenthusiastic when I say that manpower is tight so far as the armed forces go. Some time not later than a year from now approximately—and I am afraid somewhat sooner—we are going to be in a place where we are going to have to look around to find more men. When we begin to look is when we get into your business, because there are some of the things that we ought to protect, and there are some of the things that we ought to take note of, that involve the making of some very difficult decisions. They are difficult decisions because there is a great deal of difference of opinion as to who should be deferred to carry on the work of the world.

I am impressed by the representatives of different activities which believe that their activity should be supported by deferment. Some of them even believe their particular activity should be completely supported. One way to keep an individual doing something is to defer him. It won't keep everybody stable, but it does tend to encourage people to stay in place, because some of them would just as soon be there as some of the places they might be sent if they were not deferred. There is no question about the fact that there are just literally thousands of activities that are very much worth while.

It isn't letting you in on a secret exactly, but I went to college at one time. I majored, among other things, in education. They used to teach us when we were studying curriculum that it was not enough to prove that a subject was of value. It could not

go into the curriculum just because it was of value. It had to be of greater value than anything else either in the curriculum or a candidate for entrance. There is a limit to what can be put into a curriculum and, therefore, anything you put in has to be on a priority basis. Well, there is a limit to manpower. When you are making calls, you leave the veterans out, and there are certainly many reasons why you should. If you leave the fathers out—and in that case there are at least two reasons probably—if he is a multiple father there may be several other reasons why he should stay behind. There is no question about the fact that the armed forces should not take people they cannot use. There is no question about the fact that scientific and professional endeavor should be encouraged, because obviously we do not want to grow up to be a nation of nitwits and have all of these other peoples in the world outsmarting us by going to college and knowing the answers, while we stumble along and do not know the answers.

Of course there is also no question that apprentices in skilled trades are entitled to consideration. After all those people actually accomplish something while they are learning. I did not say that anyone else does not; I am merely pointing out that the people who are now pressuring me for those who are apprentices are saying at least that these people are partially productive. Of course I could say that they are spoiling material too, but that is not too good a defense. The trouble is, when you add all these arguments—logical as they may be—you are already in the red so far as manpower goes. It is one thing to print paper money and call it money, but what are you going to do with minus quantities in manpower? Well, you people have been testing a long time, you know we have had one factor all along, that is the individual who undoubtedly is a minus manpower quantity. But you have him and you are going to have to use him. It then becomes most necessary that when we determine what we need now, which is the least of our problems, that we try to determine what we are going to need ten years from now. That takes us into the training world, of course. There is not much use sending people to school for four or five years to solve a present shortage of manpower—they just do not move that fast—and there is not much sense of sending somebody now to become something five years from now that we won't need five years from now.

Therefore, one of the tough problems is to try to have enough foresight to know what kind of people, what kind of professions, what kind of scientists we shall need ten years from now. I think I have said, probably somebody told me, that atomic scientists were selling rather inexpensively in 1936, but even by other people's evaluations, they become quite expensive later on. Our first problem is to try to determine where

we are going and what we will need as we get there in this training business. I might say that one of the first things everybody would want to do when looking for manpower is to extend the age, and I can get a pretty good vote here agreeing that that is not the thing to do. I shall not have you vote, but I am quite sure you would support me in the contention that the solution is not to expand the age. I have always felt it we are bothered in the under 26 group by a numerous two million veterans, there is no use running to a place where there are fourteen million veterans who are above 26. If we are disturbed by seven or eight hundred thousand fathers below 26, I know where there are fifteen million above 26. As you extend the age upward you don't seem to find much.

If you are disturbed about the occasional highly skilled—whether he is professional, technical, scientific or just works with his hands whom we find below 26—there are thousands of them who at least claim to be all those things over 26, and there is some proof to justify them in their claims. Therefore, extending the age does not help in solution of the problem.

I might add one more thing; if we lose a million out of three million examined because the armed forces won't accept them under 26, what do you think the armed forces will find when they get to work on this second-handed quality that you find as we go along into the thirties and forties and fifties, and so forth? Therefore, everything points to using the manpower we have except one thing, the question of training. The conflict of those who are dealing in apprentices and those who deal with students occurs in these younger people. There is no question about the fact that this is a scientific age. There is probably some question of how much leadership we are utilizing and how much leadership we are losing by having men with qualities of leadership do very ordinary things, which they are probably frustrated in doing, and it would be better if someone who had less intelligence were engaged in those activities. All of those things happen, and we always get into the question of old men who want to turn nothing over to the junior executives until they have to, on the theory that if they turn it over they will never get it back. In that they are probably correct.

Just the same, a very serious problem in utilization is posed. One of the problems in engineering is to try to get all the people who are engineers to do engineering. So far as some are concerned, it is probably better to have them not do engineering. You should have tested those out. It may be requiring a little too much of you, but be that as it may, if we decide we must have certain people set aside for the purpose of training their capacities so they can give greater service to their country—otherwise there is certainly no excuse—you cannot build a system that is

fair and just by standing people aside, either because they have political influence, money, or even brains—alleged or real—you cannot create in a fair and just system privilege for any of those. The only excuse, the only reason, the only anything for deferring people to train is that, by so doing, capacities which seem apparent will be developed to an extent that the persons having the capacities can serve better and in positions of greater responsibility. Sometimes it is a little hard to keep that imbedded in the minds of the others, but unless it is, the public is going to have short timing with any sort of system if it becomes increasingly necessary to try to put the finger on the proper persons and not make more than 50 or 60 percent of mistakes when you attempt to train someone for higher responsibility.

As I walked as a layman into this store, the yardsticks were (1) not too plentiful and (2) not too generally accepted. Some yardsticks looked pretty good, but when you picked them up they seemed to move in any direction you wanted. It is a little difficult to measure when the yardstick is one that depends entirely on the fellow who is handling it. Some of the yardsticks that seemed very good to me were not too generally accepted by the public. Therefore, as with the problem of education, there was the problem of being right at least as often as being wrong, and some of the yardsticks of necessity would require several years to try to get the public to understand that we did a reasonably good job. That is one of the reasons why we are trying to lean on the fraternity—sorority, brotherhood, sisterhood, whatever we may call it—which you people here represent. Oh, we have the standing in class. When we told him something, and then later asked him, we graded on how much he either remembered or made believe he did. We use it two ways. First, in relationship between them, and then we draw a line somewhere and say, "Below this no sheepskins." That is one yardstick. Your field embraces another.

I said, probably in jest, that I did not necessarily want to swallow either one of them, but I thought by the use of both we either would compound the errors of both, or perhaps have them offset each other's errors a little bit. It is a new field, a field that you know better than I do.

With manpower as tight as it is, it seems to me that we must do the very best job we can in this setting of men aside for training and that we make no errors. I think we have done something in the last few months in getting more people aware of it. I am quite sure there is more respect for 70, just 70, than there was at one time. I am surprised by the telephone calls I get about some very outstanding student. The only difficulty was he only got 65 and he isn't even in the lower quarter of the class as a senior. Of course he is a good

man, but he is just one of those fellows who probably works a little bit outside of the normal line of people. I don't know what the end will be, but I do know in this frontier of the content of the human being, especially in the relationship to other human beings, there is very little explored. You are still pioneers, and I certainly wish you every good wish that I may because our ability to understand ourselves is going to depend a great deal on whether we have enough sense to have all these playthings we have around now in the mechanical world and not blow ourselves up before we realize the fullest benefits from them.

I don't think there is any question that the greater manpower lies in what we do not know about ourselves, and here we have been around with ourselves about as long as we have been anywhere. I am very much interested in everything constructive that you do. I came from the field artillery, and the first thing they tried to teach me was what they could do, and the next thing they tried to teach me was what they could not do. One of the graveyards of hopes inters people who had a machine that would do one thing very well and who then tried to cut hay with it or who tried to put it to some other use it wasn't made for. You see this fellow, the swimmer, whom our mutual friend told you about, made a mistake when he dove into the wheat. I now happen to be in between the testing on two sides. I have first, our own testing, with which I have been exceptionally well pleased. That is by offering something that the applicant jumps for and I think he will jump just about as high as he can.

Now, one of the fields that you are probably going much in—but frankly you will either have to do more or we shall have to have better use of what you have already done—is finding an answer to the question of how to test a man who doesn't want to play. If a man doesn't want to play baseball, how do you find out that he is as good as the center fielder of the Giants? That has come to my attention. How do you determine what a man has when he doesn't put it in the window? In fact, he has it hidden down in the cellar. It is not only a question of whether you can if you have him a week or have him an hour; one of the difficulties we get up against comes when they are streaming through the induction center and someone is visited, as they were in World War II, by a psychiatrist for 25 seconds. How do we find out all about those things? I doubt if we do. I think we would have a little more sense if we would let the lay boys sort out the tough ones, instead of trying to see all of them. It is pretty hard sometimes to sell some technicians on the fact that if they let go one possible patient they won't lose half a dozen. One of the difficulties we had with psychiatrists during the war was that they wanted to see everybody. They were not content to look at the screw-

balls that some of us laymen picked out for them and said, "'Look these over." Their percentage would be a little tough when all they got were the tough ones and there is always the chance that a person will look almost like a human being and not be but, if they are getting by, why don't we wait?

When I looked at a report this morning, (not the latest) I found there were 283 thousand potential soldiers, sailors, or airmen who are now 4-F because of a test. It wasn't physical and it wasn't emotional. It happened to be supposedly the kind that you solve with what is known colloquially as intelligence. Somewhere around 283 thousand have failed.

Undoubtedly some percentage of them should have failed, but we certainly must have something we can use rather readily a lot faster than we do now. I know very well there are a lot of people of the kind mentioned a while ago who graduated in the upper half of high school class yet couldn't pass the AFQT. When the incentive is gone it seemed to be a little difficult to use the same measure apparently established for use when the incentive was there. Perhaps the answer is, don't dive into the wheat, wait for a lake. I don't know. But I do know that we are going to approach every frontier in the world in trying to determine what the human being is and what his capacities are and how much we ought to bet on him. On the negative side, my experience would indicate an area where we have not yet learned the answers, or if we have, we haven't gotten the kind of administrative process to enable intelligent use of what is available.

It has been a great pleasure. I have always envied those who dealt in the realm of trying to find out about people, and I always hope that I shall stay at enough distance that I can consider you as miraculous and stand in awe, rather than to get close enough to be too conscious of the fact that you, too, are human, I hope.

MR. CHAUNCEY: Thank you ever so much, General Hershey. I know that everyone here deeply appreciates the fact that you came all the way from Washington today for this meeting and I know that they found your remarks both profound and interesting and also entertaining. I can assure you too, that those here will be attacking the frontiers that you outlined and hope to be able to provide instruments over the years to come that will be of greater help to you and to others. Thank you very much.

PANEL II

The Development of Useful Tests for the Measurement of Non-Intellectual Functions

The Development of Useful Tests for the Measurement of Non-Intellectual Functions

DONALD W. MACKINNON

TESTS FOR THE MEASUREMENT OF PERSONAL EFFECTIVENESS

By way of introduction I should like to say a few words about the setting in which and the purposes for which the tests that I am about to describe were developed. I feel the necessity, however, of pointing out at the very beginning that these tests are still far from fully developed and, in most cases, to call them useful is more an expression of hope than a statement of fact.

Two years ago, through a grant from the Rockefeller Foundation, an Institute of Personality Assessment and Research was established on the Berkeley campus of the University of California. The research of the Institute is directed to the discovery of the determinants and characteristics of personal effectiveness in our society, with special emphasis upon the effectiveness of persons in the arts, sciences, and professions. The subjects of our first two years' investigations have been male University graduate students about to receive their advanced degrees either in the natural or social sciences, or in medicine. Most intensively studied have been 80 graduate students drawn from 14 teaching departments. In the spring of 1950, 40 of these students were assessed, 10 at a time in weekend assessment programs running from Friday afternoon to Sunday after lunch. The other 40 were similarly assessed over weekends of the following fall and winter.

Criteria data have consisted of ratings given each subject by at least three of his instructors on the following variables: P, the candidate's Potential success in his chosen field; O, his Originality as scholar or scientist; and S, his Soundness as a person. These ratings were given prior to assessment but were not made known to the Institute's staff until after the last assessment had been completed and each staff member had rated each subject on forty variables, including the criterion variables.

In addition to several tests of intellectual functions our program of assessment has included a variety of procedures which are germane to the topic of this panel.

Increasingly in recent years has the individual's perception of himself been urged as *the* key to an under-

standing of his behavior. To tap this dimension of personality, Harrison G. Gough has prepared an Adjective Check List, which, in use, has proved to be as ingenious as it is simple. Presented with a list of 284 adjectives arranged alphabetically, the subject is instructed to check those which he considers descriptive of himself. He checks as many or as few as he wishes.

The method of analysis has involved determination of those adjectives which *both* in the first (spring) sample *and* in the second (fall) sample have shown a significant differentiation between high and low subjects in respect of any rated or measured variable, e. g., Potential success, Originality, Soundness, Spatial Aptitude, Likeableness, etc.

At the end of every assessment period each member of the staff has checked an Adjective List for each of the assessees. Thus we have for each of our subjects not only a picture adjectively expressed of that aspect of his self-perception which he is willing to make public but also a picture of his stimulus value. Applying the same sort of analysis to the adjectives checked by the staff we can discover our common perception of subjects who, in fact, are high or low on any measured or rated variable.

In addition, it is possible to derive a large number of rational scores from the adjectives checked either by the subjects for themselves, or by the staff for them, or from some combination of self and staff checks. For example, if the adjectives are classified as favorable, unfavorable, or neutral, a subject's checks can be analyzed to determine the relative favorableness or unfavorableness of his self-percept. Or an index of a subject's likeableness, as judged by the staff, can be derived from the ratio of the number of favorable adjectives to the number of both favorable and unfavorable adjectives checked on the staff's composite list for the subject, and so on.

It is not the findings which we have obtained through the use of the Adjective Check List which I wish to stress, but rather its simplicity as a personality test and the promise which it holds.

In the attempt to develop tests of significant non-intellectual functions attention has long been given to measures of opinions, attitudes, sentiments, and beliefs. Working in this tradition we, too, collected a large pool of statements to be answered by the subjects as true or false for themselves, which on empirical, intuitive, or theoretical grounds were considered to hold promise for the prediction of graduate achievement. The predictive efficiency of these face-relevant items was determined by comparing the responses of the more highly and less highly rated students in our first sample and repeating this procedure with our second sample.

An examination of the surviving items on the IPAR scale and the differentiating responses of the more highly and less highly rated subjects

to them conveys the impression that the two groups differ in respect of at least the four following dimensions:

Tempo. The high-ranked subjects seem to lead a more consistent, stabilized existence. The tempo itself may be fast or slow, but it tends to show lesser variations. The tempo seems to be "internally determined." The low-ranked subjects show more variability. They are more influenced by group demands; they are more concerned with problems of pacing, and the tempo seems to be more "externally determined."

Self-confidence. The high-ranked subjects are more self-confident. They seem to have a more stable sense of trust in the world, and are basically more optimistic. Some of this may be a reaction to their present success, but it appears to run deeper than this alone. The low-ranked subjects are more influenced by the vicissitudes of the day. Their mood depends on current achievements; they feel that luck plays a large part in success, and that good effort alone deserves to be rewarded. They are less secure, and seek assurance in fate, endeavor, etc. They are more easily swayed by other's opinions.

Time-perspective. The high-ranked subjects tend to project their goals into the future; they look ahead; they are concerned with the future use of present training. The low-ranked subjects are more concerned with immediate issues; their "learning set" is more for a present examination than for some application ten years hence.

Criticality. Both high and low-ranked subjects criticize their graduate training, but on different grounds. The Highs feel that departmental standards are often too lax, that training is not thorough enough, etc. They tend to be problem-centered. The Lows, on the other hand, complain of disagreement among faculty members, of poor teaching, etc. They tend to be more personality-centered.

Promising as this scale appears, let me add the caution that in its present stage it can be considered as little more than a set of guides and suggestions for further, more systematic research.

In planning our researches it seemed not unlikely that differences in personal effectiveness might be related to differences in esthetic judgment and preference. Specifically we wondered whether the esthetic preferences of our Highs were different from those of our Lows, and whether such differences bore any systematic relationship to expert artistic judgment. What was needed was a test consisting of stimulus material about which esthetic judgments could be made, and on which experts in art would differ from non-experts in their preferences. The Welsh Figure Preference Test (1) seemed suitable for this purpose, and Frank Barron, using some of Welsh's original stand-

ardization data plus two new samples of artists, developed an Art Scale. The 400 figures of the test (ruled and freehand figures drawn on 3" x 5" cards) were given to 37 artists and art students with instructions to sort them into two groups, those "Liked" and those "Disliked." Their judgments of each figure were then compared with the judgments of 150 non-artists (75 women and 75 men) differing widely in age, education, occupation, and geographical location.

In this manner the 65-item Barron-Welsh Art Scale (2) was derived, consisting of 40 items disliked by artists significantly more often than by people in general (p less than .01), and 25 items liked significantly more often by artists (p less than .05). It is worthy of note that the figures *disliked* significantly more often by artists were generally simple, symmetrical, and rather obviously balanced, while the figures *liked* significantly more often by artists were highly complex, asymmetrical, and rather restless and moving in their general effect. The items were made to constitute a scale, a high score on which indicates artistic preferences.[1]

When the 65 items of the Art Scale, abstracted now from the total Figure Preference Test, was administered to the assessees in our first sample, scores on it were found to be bimodally distributed; so much so, in fact, that it was clear that two distinct groups were thus defined. When the four middle-most cases of the distribution were thrown out, there was an interval of 20 points on the 65-unit scale which was not occupied by any case, 18 of the cases falling on each side of this interval.

That the figure preferences are not unrelated to other aspects of personality is revealed by an analysis of the adjectives checked by the subjects in these two groups as descriptive of themselves. These are the adjectives, differentiating at the .05 level of confidence, those who prefer simple, symmetrical figures from those who prefer complex and asymmetrical drawings.

Simple, Symmetrical
contented
gentle
conservative

1 "When applied back to the original standardization groups, this 65-item scale separated the artists from the non-artists quite effectively. The mean score of the non-artists was 16.9, that of the artists 40.25. The critical ratio was 8.46, p less than .0001. Four of the 37 artists scored below the mean of the non-artists, while four of the 150 non-artists scored above the mean of the artists.

"The scale was then tested on two new groups, 30 artists and 30 non-artists, and again discriminated the groups effectively. The means of the cross-validation samples were 18.37 for the non-artists, and 39.07 for the artists. These values do not differ significantly from the comparable values in the original standardization groups. They do, however, differ significantly from one another (C.R. of 3.97, p less than .001). Four of the 30 artists scored below the mean of the 30 non-artists, while five of the non-artists scored above the mean of the artists. Eight artists and 22 non-artists scored below the mean of the total distribution. In a new sample of 80 non-artists, the reliability of the scale, as determined by correlating the odd-numbered items with the even-numbered, proved to be .96." (2)

unaffected
patient
peaceable

Complex, Asymmetrical

gloomy
loud
unstable
bitter
cool
dissatisfied
pessimistic
emotional
irritable
pleasure-seeking

A further difference between those who prefer the simple-symmetrical and those who favor the complex-asymmetrical was revealed in a Preferences for Paintings test which was also developed and administered by Frank Barron. (3)

In this test the subject is presented with 105 postcard size reproductions in color of paintings by a large number of European artists. The paintings differ widely in respect of time and place of origin, style, and subject matter. The subject is shown the pictures one at a time and is asked to indicate the degree of his liking for each by placing it in one of four categories, ranging from "Like Very Much" to "Like Least of All." The subject is asked to place twice as many pictures in each of the middle categories as in each of the extremes, thus approximating a normal distribution.

An item analysis of the Preferences for Paintings reveals that those who like the simple and symmetrical figures like best of all portraits, religious scenes, and landscapes, and like least of all abstractions, the radically experimental, the "unnatural," and the frankly sensual in paintings.

It is the other way around for those who on the Art Scale choose the complex and asymmetrical. Among paintings they prefer the products of various modern movements in art—Primitivism, Expressionism, Impressionism, Cubism—the radically experimental, the abstract, the primitive, and pictures which portray the common-place and the sensual. They dislike religious themes in paintings, portraits of lords and ladies, and the simple representational.

Considering the striking congruences thus demonstrated to exist between figure preferences, preferences in painting, and the self-percept as revealed in self-checked adjectives, it looks as though the Art Scale is something more than a device for the measurement of artistic discrimination. It would appear to be an instrument capable of revealing the type of perceptual preferences which characterizes an individual, the one type preferring "what is stable, regular, balanced, predictable, clear-cut, traditional and following some general abstract principle, which in human affairs is personified as authority," (3) the other type favoring in perception "what is unstable, asymmetrical, unbalanced, whimsical, rebellious against tradition, and at times seem-

ingly irrational, disordered and chaotic." (3)

Though the possession of one or the other of these types of perceptual preference would certainly have many consequences, it is interesting and reassuring to note that, in our investigation, the two types were observed about equally frequently among those rated high and those rated low on potential success. There is reason to believe, as Barron has pointed out, that, "Either of these alternative perceptual decisions may be associated with a high degree of personal effectiveness. It is as though there is an effective and an ineffective aspect of each alternative." (2) It is worthy of note, however, that a preference for complexity and asymmetry is related to originality, the correlation of Art Scale scores with departmental ratings of Originality, being +.30 in the spring sample, +.44 in the fall sample.

Special attention has been given in our assessment program to the possible development of relatively simple and quickly administered tests of perceptual and cognitive abilities. Previous work in this area has suggested that simple tests of this sort may turn out to be surprisingly good measures of more complex aspects of personality. It is our hope that some of them may prove to be easily obtained and precise measures of certain aspects of what the psychoanalysts have called "reality testing." If these simpler measures can be shown to be systematically related to the more complex functions of personality revealed by more molar assessment procedures a significant extension of quantitative method to personality testing will have been achieved.

For this part of our program Dr. Richard S. Crutchfield adapted or developed a series of 14 perceptual, cognitive, and intellective tasks, performances on which have been so scored as to yield 52 measures.

Briefly noted, these tasks are:

1. *Insight puzzles* — 9 problems which can be solved through sudden insight.
2. *Gottschaldt figures* — each of which requires the subject to perceive a simple geometrical figure which is imbedded in a more complex figure.
3. *Masked figure* — a single drawing which is constructed of a word and its mirror image from below so joined that the word tends to be masked by the larger symmetrical figure. The task is to perceive the word.
4. *Tilted square* — a procedure which requires the subject to locate the vertical axis when in a completely dark room, the only visual cue being a luminous tilted square, which provides a distorting frame of reference. (from Asch and Witkien).
5. *Autokinetic movement* — with measurements of the extent, di-

rection and character of the perceived movement.

6. *Visual-motor organization* — which requires the subject to trace a luminous outline square which is exposed in an otherwise dark room.
7. *Line movement* — in which a pattern of 45 degree angle lines drawn on a roll of paper moving downward at a constant rate, is exposed behind a square aperture. The elapsed time before the perception of vertical movement gives way to horizontal movement is measured, after which fluctuations of vertical and horizontal movement are recorded.
8. *Weight-judgment* — which measures the character of shift in "adaption level" of a series of lifted weights as a function of systematic changes introduced, without the subject's knowledge, into the range of weights being judged.
9. *Kinesthetic after-effect* — a test which measures distortions in kinesthetic perception under several conditions as a consequence of previous kinesthetic stimulation, and the persistence of such distortions.
10. *Tapping* — which consists simply in determining the tempo of tapping under instructions to tap at a "natural" or "neutral" rate.
11. *Retinal rivalry* — the determination of the rate of alternation when rivaling retinal stimuli are exposed in a stereoscope.
12. *Size of constancy* — requiring judgments of equality of a variable distant triangle with a nearby standard one.
13. *Multiple choice* — employing a modified Yerkes multiple-choice apparatus on which, after three problems solvable by the same general method have been presented, a fourth problem, requiring a new type of solution, is presented.
14. *Street Gestalt Test* — in which the subject is presented with 25 incomplete pictures of objects, his task being to infer, in each case, the correct identity of the object.

It is too early to say which, if any, of these p-c tests either individually or in a brief battery will prove to be effective predictors of the criteria of personal effectiveness which we have been studying.

In our first sample, the standings of the two criterion groups, Highs and Lows in average departmental ratings of P O S, on each of the 52 p-c measures were inspected for significant differences. Sixteen were chosen as most differentiating. Each assessee was then scored on this battery of 16 p-c measures by averaging his T-scores. As would be expected, the summary scores thus obtained correlated quite highly in our first sample (+.80 with P), but the correlation showed considerable

shrinkage on cross-validation. But that these tests are measuring significant aspects of personality is suggested by an item analysis of the adjectives checked by the staff for those subjects scoring high vs. those subjects scoring low on these tests.

Adjectives checked significantly more often as descriptive of the high scorers on the p-c measures were:

1. adaptable
2. appreciative
3. calm
4. capable
5. clear-thinking
6. conscientious
7. cooperative
8. curious
9. fair-minded
10. foresighted
11. frank
12. friendly
13. helpful
14. honest
15. interests wide
16. mature
17. moderate
18. organized
19. persistent
20. poised
21. progressive
22. quick
23. rational
24. realistic
25. reliable
26. responsible
27. resourceful
28. self-confident
29. serious
30. sincere
31. sympathetic
32. thoughtful
33. tolerant
34. unassuming

In striking contrast to these are the adjectives checked significantly more often for those poor on the p-c tests:

1. affected
2. confused
3. dependent
4. dissatisfied
5. dominant
6. emotional
7. fearful
8. fussy
9. immature
10. meek
11. natural
12. resentful
13. self-centered
14. self-punishing
15. submissive
16. talkative
17. unstable
18. weak

In his factorial studies of perception and visual thinking, Thurstone (3, 4) has identified two closure factors. The Street Gestalt is a test of the first closure factor, the Gottschaldt figures a test of the second closure factor. Both of these factors, Thurstone reports, seem to transcend to some extent the visual modality. It has been his impression that those who are high on the first closure factor (Street Gestalt) tend to be outgoing and extroverted, while those who are high on the second closure factor (Gottschaldt Figures) tend to turn inward in a more introverted manner.

We find some confirmation of this in the adjectives checked significantly more often by us for high scorers on the Street Gestalt vs. high scorers on the Gottschaldt Figures.

Closure Factor I — Street Gestalt

Highs	*Lows*
adaptable	aloof
assertive	awkward
cheerful	effeminate
energetic	hostile
enthusiastic	idealistic
frank	sarcastic
good-natured	sentimental
mature	shy
natural	thoughtful
practical	
responsible	
sincere	
sociable	

Closure Factor II—
Gottschaldt Figures

Highs	*Lows*
clear-thinking	anxious
enterprising	awkward
foresighted	submissive
honest	
organized	
rational	
thorough	

The tests that I have described for you, as well as many others which there has not been time to discuss, are, of course, fairly crude. Our work with them is still rather preliminary and our findings to date are highly tentative. It is, however, I feel confident, along such lines as we are exploring that eventually useful tests for the measurement of non-intellectual functions will be developed.

REFERENCES

1. Welsh, G. S. *A projective figure-preference test for diagnosis of psychopathology: 1. A preliminary investigation.* Unpublished Ph.D, thesis, University of Minnesota, 1949.
2. Barron, F. and Welsh, G., Artistic perception as a possible factor in personality style: Its measurement by a figure preference scale. *J. Psychol.*, (In press)
3. Barron, F. Personality style and perceptual choice. *J. Personality* (In press)
4. Thurstone, L. L. *A factorial study of perception.* Chicago: The University of Chicago Press, 1944.
5. Thurstone, L. L. *Some primary abilities in visual thinking.* The Psychometric Laboratory. The University of Chicago. No. 59, August, 1950.

The Development of Useful Tests for the Measurement of Non-Intellectual Functions

RAYMOND B. CATTELL

PERSONALITY STRUCTURE & PERSONALITY MEASUREMENT

There is a species of educational psychologist, probably still represented at this meeting, to whom it will seem strange that in my title I have given personality structure priority over personality measurement. Indeed, if he goes along with the kind of personality tests most popular in schools, industry and guidance work, he will wonder why the organization of personality has to come in *at all.* If I am correct, there is something profoundly wrong in the current approach to personality testing and I beg your leave to spend five minutes analyzing what is wrong before my ten minutes of saying what I think is right; for we have a problem of re-education here.

It may be ungracious of me, since this meeting is under the auspices of educational psychologists, to suggest, as I have just done that the first thing wrong is the system of habits in the typical educational psychologist. But, in the first place the educational psychologist is unduly obsessed with items, item analysis and scaling, so that he suffers from delusions that personality is paper. If any personality test I am designing falls into itemizable and atomizable form that is generally the sheerest accident. Its numerical score is indeed more likely to be the frequency at which flicker-fusion appears, or the magnitude of a psycho-galvanic reflex response, or the ratio of certain reaction times, or oscillation in verbal output and so on. So long as personality testing is conceived only within the cramped perspective of juggling with what are essentially fragments of scholastic examinations called "items" it remains an unimaginative and trivial creature, unweaned from attainment and ability tests and psychologically impotent.

Some good personality tests will, of course, be in item form: I am only criticizing those whose whole horizon is bounded by the jargon of item methodologies. In for example, Guttman's discussions on tests, an impressive and intricate superstructure of scale analysis principles is built upon a recipe which begins "Take a number of items with the same content." Now it may be safe arbitrarily to choose content in an

achievement test in Geography, or even in an attitude test on Republicanism, but when we move from classroom material to personality manifestations this practice becomes very misleading. A personality trait can manifest itself through practically *any* content. Consequently I don't think a serious student of personality can get very interested in a complicated game of scaling procedures wherein meticulous objectivity at one point is mixed with the most naive subjectivity at another.

Becoming disillusioned by this monotonous dripping of pencil and paper items some psychologists have turned for salvation to the crystal ball, as exemplified by the Rorschach, Zondi and similar paraphernalia of the more occult clinician. These I have called "patent medicine tests" because whatever may be found out about their validity no light is thrown thereby on generalizable scientific processes and laws. For they do not deal with distinct, known psychological processes e.g. rigidity, fluency, relation eduction, corresponding to separate dimensions of personality but claim to reveal the whole of personality through a single test capsule. How hollow this claim is can be seen from Thurstone's finding that the Rorschach measures touched only one out of about a dozen factors found in perceptual responses, or from the extensive published and still more extensive unpublished literature of projective confusion and invalidity.

The remedy for the educationist's itemitis is therefore not the patent medicines of the less scientific clinician. Indeed, although the hocus-pocus of the patent medicines may seem superficially to lie at the opposite pole from the honest, antlike industry of the itemizers, the differences are of less importance than the similarities story. What they have primarily in common is a complete disregard of the necessity for discovering a great deal about personality structure before beginning to put a test together. This whole stage of mental test practice, now fortunately near dissolution, has been a strange farrago of expediency, superstition, misplaced accuracy and unjustified appeal to "pragmatic proof." Indeed it reminds one of the second verse in Genesis where "the earth was without form and void; and darkness dwelt upon the waters."

As to the nature of personality structure we have for our contingent guidance the more or less elaborate hunches of psychoanalysis and clinical psychology, which at least are scientifically respectable at the level of pragmatism. We have also such systematization of general observations as in McDougall's theories of sentiment structure, and so on. Brilliant though these leads have been, their methods are only capable of vaguely adumbrating the shape of things to come. I deny the separate existence of the so-called clinical method as an independent scientific approach to knowledge—though not, of course, as a method of therapy or

as a preliminary reconnoitering device. Scientific method has two hands —controlled experiment and statistical analysis of the uncontrolled. The clinician uses both of these, taking nature's accidents for his experiments and an intuitive use of the records of memory in place of statistical machines. (He is a one man scientific team, using the two *basic* methods of science with a justifiable roughness, but he must excuse us for concluding that his findings need precision confirmation).

Since there is very little of the total personality in its life situation that can be brought under brass instrument, controlled experimentation, the verification and extension of the existing hunches about personality structure will in my opinion be brought about largely by statistical methods. Particularly our hope lies in that wholistic, simultaneous structuring of a great number of variables which we call factor analysis. In research guided by these principles at the Laboratory of Personality Assessment and Group Behaviour ("LPG") of the University of Illinois, we have isolated and confirmed by repeated studies rather more than a dozen factors, labelled alphabetically A through O, and these factors make good sense psychologically, as unitary structures interacting in the integrated personality. One of them, the C factor, appears to correspond to what the analysts have been calling ego strength and offers an opportunity for more controlled and precise investigation of ego structure. Another, the H factor, is one of the two principal factors in the schizophrenic correlation cluster and appears to be largely a function of psysiological, autonomic capacity. Yet another (G) corresponds to the super-ego structure, and so we could pass to further illustrations if there were time. Each of these factors, first found in ratings and then in questionnaires response patterns, has been used finally as an hypothesis for the construction of special devices to measure it as a factor in terms of behavioral, objective tests. For this purpose we have used nearly two hundred newly invented tests, some of the type described by the previous speaker.

The conception of personality structure arrived at in this way is associated with several methodological developments and special conditions in the use of factor analysis which transcend what is commonly understood by the bare and basic mathematical process of factor analysis. These conditions have implications for the logical definition of a unitary trait and for the nature of personality prediction. Principally they include: oblique factors, the goal of simultaneous simple structure in several matrices instead of with respect to one only, the possibility of verification of factors by experimental influence, the notions of wholistic and conditional factors as determined by choice of variable modalities, and the strategic use of R—and P—techniques in combination. I have time for a

few words only on the last of these.

The theory here is that if a trait is to show functional unity in the fullest sense of that expression it must appear not only as a factor unity in terms of individual differences i.e. by R—technique, but must also show a unity of development, as discernible in an R—technique analysis of *increments*, and a unity of function fluctuation as shown by the P—technique analysis of the single individual. In P—technique, as you know, one can take the same set of variables as in R—technique and apply them every day, for perhaps a hundred days, to one individual, correlating the series thus obtained and factorizing the matrix. The evidence so far available suggests that some unitary traits appear both in R—and P—techniques while others appear only in one circumstance of variation, so that we must speak of unitary traits having differing degrees of *efficacy*. i.e. ability to retain their form. Knowledge of efficacy is of importance in what might be called the *natural history* of a factorial trait, under which we include information about how it changes with age, what its nature-nurture ratio is, what relation it has to psysiological conditions, to early childhood trauma, to vocational success, to therapeutic prognosis etc.

If we deliberately take a very catholic array of variables in factorization, under the guidance of the "personality sphere" concept, we seem to arrive at factors of all three modalities—some, like our B factor, are abilities; others like surgency-desurgency, seem to be classifiable as temperamental, physiologically-determined factors and others, like factors C and E—the latter being dominance or self assertion—are clearly dynamic traits. From the point of view of what the clinician means by personality structure, factor analysis has given far too little attention to this last class of traits. However, it can be shown that by starting with an exclusively dynamic set of variables, such as attitudes, one can arrive at factors which appear to be the long anticipated primary drives or ergs. Although the evidence is very recent I am convinced that factor analysis has demonstrated its capacity to reveal ergic structure, so that it is now practicable to select from the loading pattern objective tests designed to measure the strength of particular ergs as common source traits and thus to determine the roots of the average person's motivation in any given performance or situation.

The clinician, who has never made much use of common source traits, is more likely, however to be interested in the *unique* dynamic source traits yielded by P—technique. Indeed, the specification equation obtained through factorizing ergic "marker-variables" along with symptom measures (made from occasion to occasion) gives an analysis of the symptom into its ergic roots with a precision desired but never attained by psychoanalysis. P—technique used with dynamic variables, thus fulfills

the diagnostic, though not of course, the therapeutic intentions of psychoanalysis. This is but one illustration of my general point that clinical method is essentially an intuitive statistical method and that psychoanalysis comes of age when it becomes factor analysis. I have tried to show elsewhere that this method gives also a quantitative psychodynamic treatment of mental conflict and of integration. All this would seem to indicate that the prediction of personality reactions is going to require far more attention to dynamic variables than has yet been given, by mental testers, as well as an extensive use of P—technique concepts and methods in combination with those of the traditional R—technique. And although this has most application to clinical psychology, where those idiosyncratic traits which we call symptoms are to be measured and understood, it applies to *any* personality prediction. The combination of measures from the R— & P—systems, with the interaction of normative and ipsative standardizations which will be involved, should provide some interesting statistical problems.

It will be recognized that the use of factor measurements which I am advocating ventures to go beyond that implicit in most factor analytic research aims and involves additional assumptions needing examination. A majority of factor analytic researches seems to assume that the factors found are peculiar to the particular test area with which the researcher is preoccupied. The theory which has guided research in the Laboratory of Personality Assessment of the University of Illinois is that the principal personality factors operate in many different media and that researchers in different fields are probably dealing with the same factors in different dress. In terms of research planning this means that a careful and thorough sampling of factor "marker variables" must be carried into any specific research and that the planner must have an alert and comprehensive view of the whole of factor analysis and its findings, as well as an appreciation of personality theory in general.

In terms of applied psychology it means that we must abandon the isolationism of industrial, educational and clinical practices. For many years, I have found it necessary to argue in personality test work for the personality one and indivisible; by which I mean that the individual who enters the factory, the school, the home and the clinic is one and the same person, and if we are working on a true conception of personality structure we should be measuring the same factors and using the same tests in all of them. Only in this way moreover, can we take advantage of the accumulating knowledge about the natural history of these factors, converging from practice in several fields upon the basic structural picture.

Yet the industrial psychologist has been content to measure factors in his particular questionnaires or tests without any understanding of what

those factors are in general psychological terms, indeed, he is working, more frequently even to-day, with a mere correlation between some specific test and a criterion, not knowing how or why it works or whether it will disappear tomorrow or in another sample, for reasons beyond his interest. Again if we compare the educational psychologist and the psychoanalyst or clinical psychologist it is clear that they might be dealing with entirely different persons as far as their test batteries are concerned. For whereas the former attempts to predict college success almost entirely by an intelligence test or at least a battery of ability tests the latter analyses college or occupational adjustment without any preliminaries of mental testing. Indeed the only numerical estimate that a psychoanalyst is known to make is one of his patient's bank account!

The narrowness of the various applied psychologies, and their apparent belief that they can muddle along with local measurement practices has always seemed to me an incredible short-sightedness, but I have found that my astonishment is reciprocated by these practitioners, who in turn view my interest in general psychological principles either as impractical idealism or a harmless hobby. They like to talk of factors as "unreal abstractions" and to assume that "rule of thumb" has a monopoly on practicality. I am reminded of a motto inscribed in R. A. Fisher's laboratory: "The practical man is the man who practices the errors of his forefathers." Naturally the pragmatic proof of the effectiveness of dealing with personality measures related to personality structure, as determined by factor analytic and experimental research along the lines of present clinical hunches, can only come with increasing knowledge about the natural history of those factors, that is, with time and organized applications of factor measures. Then the predictions of the structures involved based on knowledge of the rise, decay, learning and interaction processes in structure, are bound to reach out farther and more accurately than the rule-of-thumb procedures of insightless "test accountants."

However, as a writer of test handbooks, I foresee a wave of protest arising at the point when psychologists have to admit, if only in the interests of accurate prediction, that psychometry has to begin to match the beautiful intricacy of nature itself. Most test users demand a right to be supplied with a penny-in-the-slot device which any grade school teacher or foreman can comprehend at once. People who could not safely be left to measure up the office carpet expect to be allowed to calculate a personality profile for a business executive, to understand what it means and to make momentous decisions in people's lives by only the simplest arithmetic. There are test publishers ready to oblige them.

The present approach, as it concerns test construction, means that

internal or *intrinsic* validation, in terms of correlation with a factor, must precede peripheral or *cultural* validation, in terms of regressions on such things as success in occupation, clinical prognosis or contribution to educational achievement. To produce a test with the greatest factor validity which its length and reliability will permit has thus been the aim in the tests which I have constructed for the Institute for Personality and Ability Testing. That a set of, say, sixteen independent factor measures, even though each has the comparative unreliability of a short test, should give *better* criterion predictions than a highly reliable but factorially unstructured test will be evident either from Shannon's writings on Communication Theory or from consideration of multiple correlation. In terms of communication theory each factor measure is an independent piece of information, whereas each score on inter-correlated tests is partially duplicated information. The independence of the elements in a profile constructed on distinct factors also permits the use of the pattern similarity index in predictions, matchings etc. in a way not possible with correlated measures.

In summary, the claim for measuring personality in terms of factors or source traits thus has three main justifications:

(1) A neater statistical handling of prediction problems in all their forms.

(2) A drawing together of the diverse fields of applied psychology with the dowry of accumulated knowledge in each.

(3) A correspondence with the functionally meaningful structure of nature and an interaction with basic principles of general psychology, but the most important of these is the last.

The Development of Useful Tests for the Measurement of Non-Intellectual Functions

JOHN DOLLARD

DISCUSSION OF PROFESSOR MACKINNON'S PAPER

MR. MACKINNON has given a lucid account of his activity in creating methods to discover "the determinants and characteristics of personal effectiveness in our society." It is, further, a very modest report, emphasizing the exploratory status of the work. No reasonable colleague could do other than hope that the MacKinnon group will have the motive, the strength and the means to continue and complete their project.

The positive features of this effort seem to me to be the following:

(1) *Importance of the problem.*

Good technique and ingenuity are of but little use on a piffling problem. The study and prediction of professional success are not in this category. We must cherish our talented people, identify them as early as possible and give them appropriate encouragement. He who has sat on the Admissions Committee of a Graduate Department will know the uncertainty which hangs over the selection of graduate and research personnel. The beloved record of college grades may only suggest the unimaginative grind, nicely habituated to academic authority. The results of our present selection procedures are, to me, always surprising. The sleepers may wake up when challenged at the Graduate level. Some of the collegiate playboys settle down. Some students reveal not only no aptitude for scientific theory but a genuine hatred of theoretical construction. They view it in very personal terms as hampering and limiting. Others gladly and intuitively subordinate themselves to a world which they assume to be orderly.

Miller and I have stated what we believe are two of the conditions for original scientific work: that a man have courage to resist authoritarian opinion, while being strongly subordinated to scientific method. How will one find these students and eliminate those who show the opposite trend, who are insubordinate toward method and accept traditional views with passionate alacrity? Further, it seems very likely that the training for good scientists and professional people begins in the simplest circumstances of life, and certainly long before the student is involved in higher educational procedures.

How are these characterological variables which favor or forbid the development of a scientist to be discovered? It seems to me that Mr. MacKinnon and his colleagues have addressed themselves to these, among other problems of unquestionable importance. It is a good trait in a scientist that he has the sense to intervene at a favorable point in the course of events.

(2) *Ingenious work.*

I was much struck with the freshness and ingenuity of the attack of the MacKinnon group. The variables P, O, S certainly have the merit of common sense. The Gough Adjective Check Test seems a promising device. It would greatly simplify evaluation procedures if the individual should turn out to be a fair judge of himself. In this case much of the torturing confusion surrounding the use of projective tests might be avoided. Through the use of this test the stimulus value of an individual can be compared as it seems to him and as it seems to others, for instance the assessment raters. The concepts of tempo, self-confidence, time-perspective, and criticality seem each and all sensible and useful. It is especially interesting that the self-percept ratings suggest these variables. I shall try later to outline a crude theory which would give a setting for them.

For my part, the suspicion that aesthetic judgments should be related to high or low P (potential success) is startling. It is probably due to my lack of information in this area that I cannot understand how the writers would have hit upon this notion. However startling the suggestion may be, it is evidently a fact that high Art Scale scores on this test coincide with a somewhat gloomy and carping view of the self while low scores identify more peaceable characters. One has the sense that somehow cultural variables have created a "built-in" correlation here, and the same logic applies to the painting test. It is also interesting, but mystifying, that these perceptual preferences are observed about equally frequently in those rated high and low on potential success.

Every scientist will share the hope of the authors that simple tests can be developed which are related to the results of molar assessment procedures. Again I find myself startled that high scores on the perceptual-cognitive tests should be related to the adjectives checked by assessment raters. Those making high scores are apparently assigned "good" adjectives while those making low p-c scores are assigned "bad" adjectives. Something strange and unexpected must indeed be operating here. At any one time one may be resigned to the fact of correlation but eventually one must know more about this remarkable connection.

(3) *Evidence of good method.*

This hardly needs stating in this

congregation. Knowledge of test construction, while not common, is evidently at hand among the personnel of this operation.

(4) *Murray's work.*

It pleases me greatly to see the imaginative work of Henry A. Murray continued in this project. That work is reported upon in a book by him and his associates, "The Assessment of Men." With marvelous common sense Murray assumed that those assessment procedures should have greatest validity which most nearly approximate the criterion situation. Least scope would then be allowed to unpredictable generalization of response. One has the stubborn sense that Murray's work is on the right track even though at the moment some of the conventional testimony to reliability and validity is absent. It is evident that the work progressing under MacKinnon and associates preserves the Murray impetus while improving on the improvised technique which was forced upon Murray by wartime pressures.

Negatives.

I must perforce and pro forma say a few things on the negative side. Actually all that I have to say is already covered by Dr. MacKinnon's disclaimer as to the provisional character of the work. I am fairly certain that he and his group would agree with the technical criticisms I am about make and am reciting them more or less as a necessary part of the scientific litany.

(1) The nature of the assessment procedures (that is, the Friday to Sunday observation period) should, of course, be more completely described. I should like, for instance, to know about rater reliability on the trait of "likeableness" which is mentioned in the text.

(2) No indication is given of the reliability of the ratings provided by the three instructors on the characteristics P, O, and S. One assumes that such ratings must be rather reliable or they would not be reported; or, at least, that the path to greater reliability is clearly seen by the writers.

(3) The reliability of the self-ratings in the adjective test would also have to be demonstrated. Obviously a test is of little use if an individual describes himself in one way on Monday and another on Wednesday.

(4) Similarly, how congruent are the ratings of assessment raters in regard to the adjective test?

(5) Of course the matter of the validity of the self-percept ratings must also arise. There is some reason to think that the account an individual gives of himself is a kind of "story" which he asks others to accept. It is the set of opinions about himself which make him least uncomfortable. Is it this set of opinions which is revealed by the self-percept?

In the course of therapeutic interviewing I have seen people whose self-percept was distorted in a complimentary direction. I have also, surprisingly, seen people who undervalued and underestimated their knowledge, skills and strength. These people I speak of may have been unique in the world, but I doubt it. People get little formal training in rating themselves. If the self-percept is a reflection of the evaluations and responses of other people these too are frequently distorted. Rivalry, sibling and professional, can lead to marked mis-evaluation by other people which would presumably be reflected in the own self-percept. Are the variables of tempo, self-confidence, etc., independently established or are they explanations of the ratings which people give themselves on the adjective test? If they are explanations, they really have the status of hypotheses and should lead to further empirical testing.

(6) Concerning the relationship between the two art[illegible]les and the two types of tempo[illegible]t which can be referred to, for my purpose, as "gloomy" and "gentle," I have the naive feeling that somehow these two scales are measuring the same or a series of connected variables. Somehow there ought to be a need for rational explanation which is not provided in Dr. MacKinnon's paper. However I well realize that a correlation is a fact of nature and when used with discrimination can be serviceable despite its lack of rational foundation.

(7) Concerning the Crutchfield p-c (perceptual cognitive) test which [illegible]uding as you remember, puzzles, weight judgment, tapping, Street Gestalt, etc.[illegible] share with the authors the hope that some such simple test will be found to be related to variables identified by more complex methods of assessment. However, quantitative hope is less than theirs. I have liked to know how shrinkage there was in the validation procedure when [illegible] of these tests were correlated with P, O, and S ratings. [illegible] actually exists, it was not [illegible] even in the preliminary [illegible] at hand.

(8) The problem of the validation of the p-c measure is, of course, a crucial one. It is suggested in Dr. MacKinnon's paper that the researchers hope that the test will be significantly related to adjectives checked by the assessment raters. Let us suppose that this relationship were high. There would still be the problem of how well assessment ratings were related to actual later "success."

A similar problem exists with regard to the P, O, and S ratings made by academic instructors of graduate student subjects. Validation in this case would have to consist of evidence of real life success,

scientific originality and personal stability over time. No evidence is offered concerning the ultimate validity of either the assessment raters' ratings or the instructors' ratings.

Answer to myself

I repeat that it is my belief, without consultation with Dr. MacKinnon or his colleagues, that they have been saying to themselves more or less the same things that I have just been saying. If I tried to answer the foregoing criticism, I would say something like this: Let these fellows alone. Give them time. They are at the beginning of a difficult task whose difficulty they recognize. They understand scientific procedures and are carefully and loyally following them. Don't ask them to face problems which they have not come to in the orderly course of events. Accentuate the positive, for there is much of it. Let your discussion have the realism of a colleague who actually faces the difficulties of original research as well as recites them. Don't pile the burden of criticism so high that the researchers can do nothing but retire to Berkeley and take up beachcombing or Russian roulette. In short, remember your own trials as a researcher and have the grace to lead a cheer for MacKinnon and Company.

I have decided to accept this advice, omitting, for the moment only, the cheer.

Can better theory suggest tests more likely to succeed?

The making of good test instruments is a laborious task. It should have every help available. Common sense is a great asset if the researcher can produce it. There is only one thing better. That is a general theory of human behavior and relations.

What I have to say here is therefore more of a comment on test-making in general than it is on the specific work of the MacKinnon group. It seems to me that our creation of significant tests is hampered to a degree we do not fully realize by the disorderly conditions prevailing in behavior theory generally. Very often the test maker is a specialist on the technical operations involved in designing a test but is lacking in sophistication about personality development and human and social behavior.

For example, I get the sense that the tester is frequently found walking warily around his subject, seeing only the external aspects of his behavior. He may try, as it were, to study man without letting him talk lest somehow the exercise of his chief human quality should bemuse or delude the researcher. The practical implication of what I am saying is that some use should certainly be found for the interview in any kind of assessment procedure which deals with the higher mental life of man. Man has the capacity not only to report on himself, to give his preferred view of himself, but also to show himself acting and thinking and thus to provide invaluable data for assessment.

Very likely the further development of the Murray procedures will exploit this capacity to the fullest.

However, I am not speaking for mere sophistication, although such is certainly better than distance and lack of involvement in human affairs.

There is every reason to believe that human behavior is lawful and orderly. If it were not, the existence of a society would be impossible. We have as yet only a dim inkling of what these laws may be or how this orderliness, intuitively perceived, is arranged.

As you know, Neal E. Miller and I have been working with the laws of learning as a form of systematic explanation in this field. Our reasons for being hopeful about learning science are several: The first is that there is a large body of experimental evidence on learning and some over-all agreement as to theory. For example, practically every experimenter of consequence believes that habits can be learned under conditions of drive and reinforcement; differences might exist as to whether these are the only conditions under which learning can occur. However, so far as the general public of science is concerned, there is a considerable body of theory on which there is technical agreement, as well as a smaller body of theory which is still under experimental debate.

The second great consideration which makes us look hopefully at learning principles is a type of cross-cultural analysis. We know that from the physiological and anatomical side the differences between men of any race or culture are minuscule. Their bodies are rigged and wired in substantially the same way. Apparently they learn in the same way. Fundamental drives are at least more or less similar. The differences in skin color, hair form, and the like, which are so highly and dangerously cherished in the English-speaking world, are unimportant from the physiological and anatomic standpoint. One may say, therefore, that the men with which culture begins are the same.

But looking at the scene of adult behavior, one finds an astonishing variability. The three thousand different societies which have been identified by anthropologists produce three thousand different kinds of individuals, characterized by different food appetites, different sex habits, different economic practices, different senses of guilt, different objects of fear, and a multitude of other differing traits.

How can it be, then, that from a single kind of man so many different patterns of adult habits are produced? Our answer to this question is that these habits must be learned. Research has shown, furthermore, that there are a few powerful principles which account for such diverse learning from the single substrate. It is this consideration which especially focused our attention on the learning field in Psychology and which makes us hopeful that a better

understanding of learning processes will lead at one and the same time to a better general understanding of human personality and social relations.

We believe, as a result, that the tortured testers, seeking to get a laboratory grip on the slippery eel of human personality might do well to learn thoroughly what is known about human learning and to seek, in these principles of learning, stimuli which will indicate favorable variables for testing or favorable situations for evoking such variables.

Since we have not set ourselves the task of creating such measures ourselves we cannot be certain that the use of learning principles would be an economy in devising tests. It hardly seems credible, however, that they would not be more serviceable than intuition or common sense. Principles which appear with so much order and coherence in rats, dogs, cats, monkeys, chimps, children, college students and Ph.D.'s could not fail to be suggestive in the measurement field.

An attempt at application.

I would not have it supposed that I am any less willing to stick my neck out in a venturesome project than is Professor MacKinnon. The critic should have no less courage than the creative scientist. I will therefore make an attempt to make a brief behavioral explanation of the variables tempo, self-confidence, and time perspective. This will necessarily be an ex post facto explanation since it is Professor MacKinnon who has evolved these descriptions. I think it is possible, however, that with an adequate knowledge of learning science some of them might have been surmised before a test based on hunch had shown them to be related.

First as to tempo. This is described as a stable rhythm of activity in the students who were rated high by their instructors. Presumably their activity is less variable than the low-rating students and less responsive to immediate stimuli of discouragement. Let us remember that the tasks of the graduate student in science and medicine (though to a less degree in social science) are laid out in serial order. First units are learned first. Learning of second units is based on proper acquisition of first units, etc., very much as one must know arithmetic before one learns algebra and something about the use of equations before one learns calculus. I should guess that a student whose activity unrolled in an unruffled pattern would have a better chance of acquiring preliminary units in the right order and would thus make steadier professional progress. Students whose rhythm of activity was more variable might miss out on necessary elements in the serial pattern and thus be disadvantaged in later learning.

As to self-confidence. This trait is presumably based on past successes. The student has responded in the past, been rewarded, has learned, and this transaction has left behind a de-

posit in his character. In the Miller-Dollard lingo, we call it "hope." Hope is a self-administered reward. It can bind into place the activities of an otherwise dull and frustrating day. It can be attached to such thoughts as "There is always some bad with the good," or "Now I'm half way through," or "I've been through it all before, and it came out all right," and so forth. This type of self-reward can bind the activities of daily life into place and enable the student to bear the long years of training which are required of modern professionals. We should suppose, therefore, that self-confident or hopeful students would learn better and therefore would merit better assessments of likely success.

In regard to time-perspective. I have emphasized the length and often tedious character of the learning required in years of professional preparation. In this connection time-perspective can function to motivate the student when near stimuli do not have this effect. Possessing adequate time sentences, if you will permit the expression, enables one to use motives in the present time which are evoked by the prospect of a distant failure; and similarly, to enjoy and relax in the prospect of a remote success. The student who is correctly time-oriented is therefore much more autonomous as far as his professional learning is concerned. He is less distractible by immediate stimuli and is able to motivate and reward himself. He should, of course, learn better and get better ratings.

Naturally I am far from satisfied with these systematic suggestions. I have simply been trying to say that some of the actual findings do seem to follow fairly naturally from a learning analysis of the situation in which the professional student finds himself. I trust that the analysis is somewhat plausible because I would wish to strengthen the hope that test variables could be selected in a more orderly way if test makers had or used a better general theory of personality. I think it further to be in the interest of science to suggest that the learning and over-learning of the elements of present-day learning science might be of real advantage to the maker of tests.

The Development of Useful Tests for the Measurement of Non-Intellectual Functions

SILVAN TOMKINS

A DISCUSSION OF "PERSONALITY STRUCTURE AND PERSONALITY MEASUREMENT" OF R. B. CATTELL

CATTELL believes the educational psychologist is to unimaginative, the clinical psychologist is too imaginative, the experimental psychologist and his methods unattainable and, that since the earth is without form and void and darkness dwells upon the waters, salvation can only be found in axis rotation. I believe that the flood can and will inundate testing, psychometric and projective, and factor analysis alike, unless all of these are revitalized by theory and experiment. If it were true that educational personality tests are trivial, projective tests wild, experimentation impossible then factor analytic techniques would in fact command an inflated value on the psychological bourse. We are not faced with these options. Tests, questionnaires, inventories, projective techniques can be salvaged. Experimentation is possible and its area can be continually expanded.

Let us consider first the matter of structure versus measurement. Cattell has said "Indeed, although the hocus-pocus of the patent medicines may seem superficially to lie at the opposite pole from the honest, ant-like industry of the itemizers, the differences are of less importance than the similarities. What they have primarily in common is a complete disregard of the necessity for discovering a great deal about personality structure before beginning to put a test together." I agree entirely. It has been characteristically American to measure first and ask questions later. But we cannot blame it all on Americans. Binet was a Frenchman and Rorschach a Swiss and these two began the testing movement. Measurement has outrun theory for a simple reason. Psychometric and projective tests had their origin in and owe their continued support to the needs and demands of service agencies, educational institutions and mental hospitals and clinics. Binet was asked to predict scholastic achievement not construct a theory or a test of intelligence as such. Rorschach did not set himself the task of constructing either a test or a general theory of personality. He was primarily interested in the differential diagnosis of the mentally ill. Both testing move-

ments have straddled this issue with varying degrees of discomfiture ever since. There has been an enduring interest in the prediction of scholastic achievement and at the same time an unwillingness to relinquish the notion of the measurement of a theoretically more remote construct-"intelligence." The projective testing movement has been equally troubled. It has not wanted to divorce itself from the general theory of personality and yet feels it must be responsive to service needs—is this person a schizophrenic or a depressive patient—will he respond to therapy—to what kind of therapy will he probably best respond? These are practically important questions. We do not know whether this is the type of question the personality theorist should ask. We have misgivings when projective test results are used to bolster a classification of mental diseases that is admittedly obsolete. But the projective tester and the educational tester feels committed to answer questions whether or not he has posed the questions. Both psychometric and projective tests were ultra empirical in design. In both cases one knew what one wanted to predict and experimented with items, and ink blots until they did in fact predict. Each type of test was almost equally unembarrassed by theory. If one achieves a certain degree of success in a venture of this type then the obvious next step is to do a somewhat better job. What seems to be called for is more precision and then measurement begins to outrun theory. Why did neither movement sustain an inquiry into the nature of intelligence and personality? My guess is that this did not happen because of the success of both psychometric and projective tests. Theory breeds primarily on conceptual dissatisfaction. Let us make no mistake about it—these tests are exceedingly powerful instruments. The intelligence test yields us information which might otherwise require a few years of observation. A Thematic Apperception Test will yield information about an individual which very intimate friends may not acquire in the course of a lifetime. Those who are unfamiliar with the mental hospital and clinic are inclined to underestimate the amount and quality of feedback information which is available to the clinical psychologist. There are literally thousands of instances in which a projective test has enabled the clinician to understand and predict the course and outcome of mental disease and therapy. This feedback is neither as systematic nor as reliable as the type of information we obtain from a well designed experiment—but it is frequently as good, and sometimes better than the information derived from experimental or statistical surveys which are not so well designed. It is this very degree of predictive power of both psychometric and projective tests which I would hold partly responsible for the condition which Cattell and I deplore. It has in effect taken the urgency out of the theoretical ques-

tions—what is the nature of intelligence and personality? It would, I submit have been a better thing if both the psychometrician and projective tester had initially failed in his efforts at prediction of a criterion. It is this purely empirical, a-theoretical, criterion-bound characteristic which is responsible for the low esteem in which the whole testing movement is held by the parent science. The tests are not rejected on the grounds of low validity—they are in fact as valid as most theories are true. It is rather because even when valid they contribute so little to the general fabric of the science. The clinical testing movement is today undergoing the same type of validity soul searching as the psychometric testing movement experienced a generation or so ago. Does the Rorschach M response really mean inner-life—do the great scientists really have an unusual number of such responses—is color shock really due to the color of the plates—let us take the color out and find out. These and a host of similar questions are providing Ph.D. thesis material for our young clinicians. You may say "it's about time." I would disagree. Errors of interpretation of projective tests would in the long run harm us less than the diversion of our main energies to problems of test validation.

Although I agree with Cattell that measurement is being overemphasized in the entire testing movement I cannot escape the impression, that he bears some prejudice against projective techniques. By prejudice I refer to the state of mind which requires a certain level of confidence to believe one kind of theory and another level of confidence to believe a different kind of theory. Consider his statement "How hollow this claim is can be seen from Thurstone's finding that the Rorschach measures touched only one out of about a dozen factors found in perceptual responses." He might have cited many other studies exposing low validity and as many that were on the positive side but he chose to emphasize the negative findings and to present as negative a finding which reads in Thurstone's (6) monograph as follows:" The test was given to our subjects by Dr. W. A. Varvel, who has given considerable study to this test. A number of separate scores were determined for each subject and Dr. Varvel may publish separately an analysis of these Rorschach records. Only two of the scores were included in the factorial analyses namely the scores denoted Total Responses and Perceptual Organisation . . . the two Rorschach scores appear here by themselves in the doublet factor K. The other saturations are smaller, and we have not been able to formulate any hypotheses concerning the interpretation of this factor beyond that of the Rorschach test specific." It should be noted that less than one in ten of the conventional Rorschach scores were included in the factor analysis. It should

not be too surprising if these two scores do not appear in a variety of perceptual factors. That even these scores are not without psychological significance appears in the appendix of the same monograph, in which Thurstone, reporting a study of campus leaders writes "The Rorschach test was scored in several ways, and these scores showed interesting differentiations for the campus leaders. The leaders excelled in the total number of responses, they were very markedly superior in the perceptual organisation score and they showed greater latency for color cards. They gave a smaller number of responses to the color cards in comparison with the rest of the perceptual battery, so that these scores cannot be interpreted by the same factors as far as we can determine at present." In other words the test was effective in discriminating one group from another, but since only two scores were used in the factorial analysis and since these scores did not seem related to other perceptual factors—no more than this could be said. This would seem to me slim evidence on which to reject a technique that has been useful for a period of over thirty years. I should like to emphasize that validity is not a characteristic of a test and truth is not a characteristic of an experiment. Truth and validity are characteristics of statements we make and inferences we draw on the basis of experimental or test data. We can and do make false inferences from experimental as well as test data and we do make valid inferences on the flimsiest of evidence. We need not, because of invalid inferences, scrap either the general method or the particular experiment or test. We must indeed stop drawing incorrect inferences from a particular test when we learn that we have been making errors, but we may or may not want to stop using the test. Cattell has called projective tests "Patent medicine tests" "because whatever may be found out about their validity no light is thrown thereby on generalizable scientific processes. For they do not deal with distinct, known psychological processes e. g. rigidity, fluency, relation eduction, corresponding to separate dimensions of personality but claim to reveal the whole of personality through a single test capsule." I agree that the crux of the matter is that validity studies need not throw any light on generalizable scientific processes—in the case of both psychometric and projective tests. But they may be analyzed in such a way if one wishes to do so. There have been, for example, an increasing number of factor analyses of the Rorschach in an effort to do just what Cattell says these tests cannot do. Shelagh M. Cox (2), working under Eysenck's direction, found that the first factor to emerge from the Rorschach was a productivity or fluency factor and that there was a correlation obtained between the criterion (normal-neurotic) and the first two factors to appear in the analysis of the Rorschach categories

of .78. This fluency factor was one of four which also appeared in Whittenborn's (7) factor analysis. Cox reported "The general factor of productivity found in the present investigation links up with Petries findings. She carried out an analysis on the inter-correlation of results on tests involving the giving of as many responses as possible in a given time. Such tests included responses to a black and white Rorschach ink blot, number of round things, trees, things to eat, etc. A general factor of fluency was identified. This factor had its highest correlation in number of responses to the black and white Rorschach blot. This can be compared with the high loading of number of responses for the general factor in the present study. A similar factor was identified by Sen with Indian adults." This latter was also a Rorschach factor analytic study. The point is that there is nothing about any kind of data, or test which necessarily dictates the kind of analysis one may make of the data, or which dictates the purposes for which this analysis may be made. I would however agree that if one's interest in tests is to be theoretical, test construction should in the first instance be inspired by the theoretical interest. But you may say "that would not be test construction that would be indistinguishable from pure experimentation." That is exactly my point—that today's tests are poorly designed experiments. They answer no one question very well. They typically answer too many questions, yielding too much information to be useful in the reconstruction of what has gone into the process of responding to the test. Many tests resemble defensive camouflage. They present information in such a way that one is fairly certain not to be able to use it to answer very searching questions. Implicit in my criticism is the assumption that the test should be more than a measuring instrument of a single variable or factor. A good measuring instrument, let us say a psychological thermometer or voltmeter would I think satisfy most testers, including Cattell, who would use factor analysis to discover the basic dimensions and then devise ways of measuring the pure factor. But this is, I submit, too low a level of aspiration for the testing movement. The thermometer is a wonderful measuring instrument, but is not likely to yield a theory of heat, nor would a voltmeter yield Ohm's law. The more usual situation is one in which a theory and a lawful relationship are used as a basis for creating measuring instruments. A group of measuring instruments may, on the other hand, be necessary to discover laws—but in the personality field our problem in general is that we have too many things to measure and we do not know which variables will be the important ones, and what the important relationships between variables might be. Shall we use physical models, or models such as chemistry has found useful in organizing organic substances or what? Factor

analysis, if we have been lucky enough to include the appropriate information may tell us what are the important dimensions. But from that point on we are on our own and it is not a short distance from that point to our goal. Not infrequently however we are not altogether sure what this dimension is. It seems to be something common to many things which seem different. What is called for is an inspired guess. I would suppose that if one were capable of making that inspired guess at the end of the factor analysis, it is not inconceivable that one might have made it before.

The testing movement has taken little responsibility for conceptual models. Gulliksen (4) has made an important suggestion I believe, in his concept of intrinsic validity. I would urge that testers, psychometric and projective, take the next obvious step—drop the concept of validity altogether, and concern themselves with the truth or falsity of their ideas and assume the scientist's general responsibility for the generation and testing of important ideas, devising tests in such a way that they are indistinguishable from experiments—so that one test leads to the formation of another in the way in which one experiment leads to another experiment. I am suggesting that an experiment can not be identified by the nature of the material it employs —brass is no more appropriate for experimentation than paper and graphite. Cattell has said that scientific method has two hands—controlled experiment and statistical analysis of the uncontrolled. I agree, but if it has no brain it will be limited to scratch reflexes no matter how refined its statistics or how shiny its brass. But how can one combine theory and testing? How can one experiment with a test? I am at the moment working on a theory and a test which will test the theory. I am concerned with what strikes me as an important difference between two kinds of people both classified as psychotic. One kind of psychotic begins at an early age to experience difficulties, with progressive deterioration involving more and more of the personality. He is likely to be institutionalized and spend the remainder of his life in a mental hospital. The other type of psychotic is given to occasional and sometimes periodic outbursts of violence, depression or elation. Such a one may be in and out of mental hospitals or may be hospitalized only once or twice, but is normal between episodes. The nearest diagnostic labels are the schizophrenic and manic-depressive psychoses. Both groups seem much concerned with the problem of violence. How is it possible that the same problem could result in such radically different consequences? I entertained the hypothesis that the manic-depressive is primarily concerned about controlling his own aggression and that he feels guilty when he does express aggression or even when he feels hostile—further,

that he is able to emerge from periods of atonement and then lead a normal life, for a while. The schizophrenic, I had reason to believe, was more afraid of the physical danger associated with the expression of aggression, whether that aggression was expressed by himself or by another person. The schizophrenic remains anxious because he avoids situations which he might have learned to master. Withdrawal, the psychosis itself, is a stable adjustment because it protects the schizophrenic from the dangerous situation and discourages relearning. The manic-depressive is the instrument of his own partial cure—he punishes himself or gets himself punished and as his punishment increases his guilt decreases until he is no longer in an elated or depressed state. He is normal and free again to aggress until his guilt becomes intolerable. Such at least is my theory. One can test it by projective technique. One can present two types of situations for interpretation. In one a group of people are shown fighting—in another a person is shown bleeding from a physical injury. If the theory is correct the picture of people fighting will be distorted—either exaggerated or minimized (compared with normal controls) by both groups but that the manic-depressives will be normal in their interpretation of physical injury while the schizophrenic will exaggerate or minimize physical injury covarying with his interpretation of the fighting situation. Preliminary evidence supports the hypothesis. If a schizophrenic denies that there is any fighting going on in one situation he denies that the man's finger is bleeding in the other situation. If he exaggerates the intensity of the fighting in one situation he also exaggerates the seriousness of the physical injury. The manic-depressive distorts the fighting situation just as the schizophrenic does, but his interpretation of the situation showing physical injury does not differ from the normal interpretation. If these findings should receive further confirmation one would possess at once a test *and* a theory of the nature of the two major psychoses. One should then proceed to improve both the theory and the test—and not standardize one and experiment with the other. Admittedly one does not standardize theories—why should one either invite or reject the likelihood of obsolescence by a double standard for theory and test? Only on the assumption that a test is a measure of a single parameter, that it is the best possible measure of that parameter and that it is the best possible parameter to measure. I would suggest first that a test may measure more than one parameter, that the adequacy of the measure is a function of the adequacy of the underlying theory and what we choose to measure will depend first on what we think may turn out to be important but what we continue to measure will depend on whether these measures yield us important functional relationships. How-

ever, what we come to regard as important will depend on our level of aspiration. In short I am recommending that we regard the test as an experiment designed to confirm or reject a hypothesis rather than simply a measure of a single parameter.

Now we arrive at Cattell's third assertion—"since there is very little of the total personality in its life situation that can be brought under brass instrument, controlled experimentation, the verification and extension of the existing hunches about personality structure will in my opinion be brought about largely by statistical methods. Particularly our hope lies in that wholistic, simultaneous structuring of a great number of variables which we call factor analysis." In another context he has said "For many years I have been arguing in personality test work for the personality one and indivisible, by which I mean that the individual who enters the factory, the school, the home and the clinic is one and the same person, and if we are working on a true conception of personality structure we should be measuring the same factors and using the same tests in all of them." One should not entertain both of these propositions at once. If the person in the clinic is the same person at home and in the factory—then he is also the same person in the psychological laboratory. It is one and the same person who is studied in all of experimental psychology. It is a person who perceives, remembers, thinks and acts even though experimental psychology has only recently recognized the role of "personality" in "perception." The person not only *can* be studied in the laboratory—he has been for a hundred years. The extension of the scope of such study depends upon the extension first of theory and then of technique—there need be no real limitation in either direction. But there is something more to Cattell's misgivings about the laboratory study of personality. First, he identifies the laboratory with the brass instrument—an unnecessary limitation upon the possible paraphernalia of the well equipped personality laboratory, which may possess couches and pictures as well as brass. Second, there is the implicit assumption that nature in the raw must somehow be transported bodily into the laboratory in order to study personality. It is as though, for example, people must be observed falling in love under the voyeuristic atmosphere of the laboratory if the phenomenon of romantic love is to be studied in the laboratory. This is no more necessary for the personality theorist than it is for the theoretically minded physicist to have automobiles running into each other in his laboratory in order to study the dynamics of irresistible forces and immovable objects. This brings us to his third assumption which is quite explicit—the assumption of wholism as a preferred mode of approach in personality theory. If one means by wholism that one should leave out nothing important or rele-

vant when one studies personality I should agree, but if one identifies wholism with nature "in the raw," that in order to study personality one can exclude no fact, no phenomenon no matter how trivial, then I disagree. One must have scope in personality theory and experiment, but scope is not to be identified with the assumption that all facts were created free and equal and none should be excluded from theory construction. If all facts are not equally crucial for a theory of personality then we cannot be certain that the crucial phenomena may not be studied under laboratory conditions. I have been trying to defend the possibility and importanee of laboratory study of personality. But I would agree with Cattell that an experiment is not limited to the laboratory. Factor analysis may be used in an experimental manner as Cattell is doing—this too is profitable—depending upon the extent to which this type of experimentation yields cumulative theory. My preference is for stating the theory explicitly, at the beginning of the investigation and devising the technique most appropriate for testing the theory, rather than hoping that if one pours enough sand through a sieve the gold nuggets will be found.

I should like to close with three sets of suggestions for the personality and psychometric testing programs. First, inasmuch as both types of testing are oriented toward application, a general theory of achievement is urgently needed. If one wishes to use "intelligence ... predict ... academic achi... one should spend mor... theory of achievemen... on a theory of intelli... as intelligence is no more ... component necessary for ... t. Secondly, I should like to ... ention to a neglected research ... which owes its neglect to the restriction of vision of the personality and psychometric testers. This is the borderline inquiry of the relationships between personality and intelligence. We are presently undertaking a series of studies in the relationship between intelligence and psychophysical judgment, rigidity in thinking and general social ideology. Schall (5) has demonstrated that the more intelligent are capable of more accurate psychophysical judgment. Basescu (1) has demonstrated that the more intelligent shift more quickly than the less intelligent when reinforcement is increased in a concept formation and shifting experiment. Goldstein (3) has shown that generalized social attitudes such as prejudice are inversely related to intelligence.

Finally I want to present a set of what I hope will be helpful hints to test constructors—those general characteristics of personality which seem to me most often disregarded in test construction. First, the principle of *compensation and lability of energy*. Within certain limits the personality is organized like an army. It has reserves of energy and uses these when necessary. It is usually adept

at shifting these reserves now to one place and then another. Wherever the tester is—there will the reserves be likely to be shifted. This has made it very difficult to study the effect of stresses upon personality and intellectual function. My general recommendation is to look, as unobtrusively as possible in the rear of the front lines—in the periphery of the psychological field—or after the battle is over—in the post experimental period, during which a "debt" is usually being discharged. Secondly, the *variability* of a personality parameter is the *basic parameter*. Again and again psychologists have followed the will of the wisp of the general parameter—e. g. fluency, aggressiveness, etc., only to discover that the parameter becomes fragmented under more careful study. There are no such parameters as generalized parameters. This has had to be painfully rediscovered with each new personality characteristic studied. One must think in terms of the specific conditions under which the parameter appears—its generality and variability are the critical phenomena of study. Only a one-celled organism could be described by a general parameter or factor, that is, give essentially the same type of response to all situations. Third, the phenomenon of *gradient enhancement*. The organism is a system such that any gradient within the psychological field can be enhanced, its influence magnified if all other conditions in the field are kept reasonably constant. One cannot generalize from the situation in which one parameter has been varied and all other parameters held constant to situations other parameters do vary. Any steady state changes the organism radically, heightening the dominant gradient and homogenizing the rest of the field. Four, the phenomenon of *restriction of range* which is a special case of gradient enhancement. When one asks questions or calls for behavior under experimental or test conditions one usually places a restriction on the range of stimuli and the range of alternative responses. We cannot generalize from test or experiment unless we have a theory of the effect of such restriction of range with respect to stimuli, responses and coding categories and a self conscious variation of the restriction of range. Five, the principle of *equipotentiality and equiactuality*. The person either has alternative methods of achieving one goal, or can learn alternative methods under the press of test or experiment. Therefore one should not equate performance with the underlying mechanism. The latter can only be fathomed by systematic manipulation of aspects of the test or experiment, for example by tests of transfer built *into* the test. Different individuals may have different ways of achieving the same end. One must therefore vary critical aspects of the test or experiment until these achievements are no longer identical. These principles seem to me a sample of a very much greater population of principles which are commonplace in experimental study of

personality and which should become commonplace in test construction theory.

REFERENCES

1. BASESCU, S. Intelligence and learning. Ph.D. Thesis. Princeton U. Library.
2. COX, S. M. A factorial study of the Rorschach responses of normal and maladjusted boys. J. Genet. Psychol., 1951, *79*, 95-115.
3. GOLDSTEIN, M. Study of Prejudice. Ph.D. Thesis. Princeton U. Library.
4. GULLIKSEN, H. Intrinsic validity. Amer. Psychologist, 1950, *5*, 511-518.
5. SCHALL, H. Intelligence and Psychophysical Judgments. Ph.D. Thesis. Princeton U. Library.
6. THURSTONE, L. L. *A factorial study of perception*. Chicago: Univ. Chicago Press. 1944. P. 148.
7. WITTENBORN, J. R. A factor analysis of Rorschach scoring categories. J. of Consult. Psych., 1950, *14*, 261-267.

DISCUSSION

PARTICIPANTS:

DONALD W. MACKINNON, RAYMOND B. CATTELL, HENRY S. DYER, LLOYD G. HUMPHREYS, JOHN P. KERNAN, EDWIN G. FLEMMING.

DR. MACKINNON: I am indebted to Dr. Dollard for the thoroughness with which he reviewed my paper. I appreciate the positive things he said, but particularly I am indebted to him for the negative things, the criticisms, because, it seems to me, these criticisms are likely to be the criticisms that you felt as I was speaking to you, and these criticisms, it seems to me, arise out of a misunderstanding for which I take responsibility.

If I had met Dr. Dollard's first criticism and told him and you what happened from Friday afternoon until Sunday after lunch, many of the points which were raised would have not have been made. I restricted my remarks, and even so ran overtime, to stressing those parts of our program which seem to me to come nearest to being the development of what the panel title is, namely, "Useful Tests for the Measurement of Non-Intellectual Functions," these useful tests being tests which eventually can be presented in paper and pencil form, widely distributed, and administered as group tests.

I should like to say that I second heartily all of Dr. Dollard's emphases upon theorizing and I should like to point out that we in California spent six months doing nothing but theorizing before we saw a single subject. Indeed, we take some pride in the fact that we spent more time in thinking about how outstanding people become that way, than have many others who had ventured upon assessing or testing programs. A good deal of that theorizing, and many of our thoughts about how personal effectiveness is achieved derived from the work of Mr. Erik Erikson. Mr. Erikson was at that time a member of our staff, and much of our theory as to how the individuals we were studying became the kind of persons they were, was very congruent with the notions which Mr. Erikson subsequently presented in his book "Childhood and Society." One cannot read that book without knowing we were influenced by both its theory and its cross cultural data.

I am very sorry that I gave any impression that we operated without the interview. Not only did we have

the interview, we had four interviews. Erik Erikson interviewed each subject for an hour and a half. In this he tried to get the individual to speak about the way, as he saw it, in which his ego- and group-identities had been developed. We had an hour-and-a-half interview on the subjects philosophy and values. We had an hour interview on the life history, and a half-hour interview on the professional field of our subjects, letting them tell us why and how—as they saw it—they became interested in their chosen professional fields. I did not report these to you because of the limitations of time, and because, it seems to me, interview material (valuable as it is) is not the kind of material which I thought it was my assignment to discuss with you this afternoon.

With respect to Dr. Dollard's point that some of the dimensions—tempo and self-confidence, time perspective—which we discovered, might have been surmised prior to our testing, I should like to point out that they were surmised, and it was just because we surmised that these might be important variables of personality that we wrote items to top them. As a matter of fact, we wrote 900 items for our original IPAR questionnaire and if we had not written those items these particular dimensions would not have come out, because you do not get out of a test anything more than you put into it. We had a lot of other surmises, some of them were to me more exciting than these, which, I regret to say, did not stand up in the application of the questionnaire to both our first and second samples.

I should like also to point out that I feel we have in our assessment program made one further important step ahead which I did not have time to discuss, namely, the development of situational procedures which make it possible to control the situation so that everybody in the situation is presented with exactly the same kind of stimulation. Out of this I think we will be able to realize at last a great deal of the promise that many of us have felt existed in situational tests and procedures but which to date has not been realized because most situational procedures allow so many uncontrolled factors to operate.

We have lots of reliabilities which I should like to quote if there were time. There isn't, and again thank you, Dr. Dollard.

DR. CATTELL: I want to thank Dr. Tomkins for what seems to me a very brilliant and sound commentary. As I said, since I am strongly interested to hear also the critical reactions from the audience, I shall confine myself to commenting on Dr. Tompkin's remarks, on just two points. First, in regard to the much maligned Rorschach, I should say my objection is not to the Rorschach but to its devotees. The Rorschach is one of about two or three hundred tests which are available in the literature for clinicians to work upon, all with about the same degree of promise.

For some reason they could not lift their eyes beyond one of them. Therefore, my criticism is merely of the kind that it would be, shall we say, of rum. Rum is a good thing in itself, it is only the man who feeds on that and nothing else of whom one is entitled to get critical. However, I know that in view of the fact that Dr. Tomkins has done wonderful things with the Rorschach, more, in my opinion, than any other person, and in view of, as he says, the length of our past friendship, I do not think we should allow the Rorschach to perform one more foul deed by separating us!

The second issue is one of the precise meaning of empiricism and theorizing. I think we are in some confusion over the use of the words "hypothesis," "theory," and so on. I don't mean Dr. Tomkins and I are, but I mean the general psychological atmosphere of the present day seems to be in confusion. I am an empiricist, I am for going to the facts first and I am against sitting down and dreaming up theories from nowhere as has been very prevalent when we have had schools of psychology, each with massive, pretentious theories to cover the whole field—without having a foot down anywhere in it. The hypothesis should arise out of empirical data, but it cannot do so unless we go beyond everyday observation, that is to say we have had for two or three thousand years of culture, in the narrow sense, access to human behavior and observations on it, and I don't see that by mere continued observation of existing data on the clinical level that we can hope to get any further than Aristotle did. We have got to go out with new methods and more powerful means of analysis and then construct our theories from those.

Factor analysis along with experiment is just one way of getting new *relationships* out of which we can develop hypotheses; and so, out of the hypotheses, theories. Consequently I agree entirely with Dr. Tomkins that the emphasis is on the inspired guess *when you have got your factor.* The dozen or so factors which personality research has now produced are a dozen hypotheses, each of which could keep two or three hundred people busy for ten years. These are the really new hypotheses in the picture of personality theory. My reason for preferring factor analysis rather than controlled experiment is that you deal with several variables at once instead of just two. Often a theory or a hypothesis cannot be operationally defined adequately by one variable alone. It is a *pattern* of things, so we need to turn to either a systematic building up of the pairs, dependent and independent variables on the line Dr. Tomkins is doing or else take a more global approach now possible through the factor analysis.

Those are my two comments regarding Dr. Tomkins, but I should like to beg leave to make one com-

ment that I think Dr. MacKinnon might have made about his critic. The latter seemed to me to get away with—well, to get away with murder—as far as scientific method is concerned on one point. He said because there are not very marked hereditary differences between racial groups and since the patterns of personality and different cultures were extremely varied, that personality must be largely a learned pattern and that learning theory is therefore largely the means of understanding it. Now this formulation confuses what any statistician will recognize as the intragroup and intergroup variants. It may be true that there are no big differences between groups but there are big hereditary differences within groups and, of course, we could always put groups together for experimental purposes by selection of individuals which would give very big innate differences between them.

I stress this because I think we have neglected the innate side of personality for the last twenty years and only this last month there has appeared a paper, quite a startling paper, by Eysenck from Maudsley Mental Hospital. In this very adequate and methodologically interesting twin study in which he concludes that something like 60-70 per cent of the variance in neuroticism is determined by heredity and only 30-40 per cent by environment. (I quote from memory.) That is a very high figure, higher than anyone would have expected, and I can see no major flaw in any part of his own procedure. So, although I am in favor of using learning theory and have always explicitly recognized what I have called "environmental mold factors" which are to be regarded as patterns of change in many variables imposed by social institutions, we must also recognize with equal explicitness that the origin of others among these factors lies in the biology of the organism, and if Eysenck is right it lies in the truly genetic influences.

CHAIRMAN DYER: We haven't as much time as I thought we would have, but what time we have is now yours.

DR. HUMPHREYS: A couple of our speakers made the statement that measurement has outrun theory. I should like to suggest what I consider to be a more precise statement, that is, that the accumulation of measurement devices has outrun both theory and data.

MR. KERNAN: I have two very similiar questions that I should like to ask Dr. MacKinnon. First, is it not possible that significant differences found in the preferences of artists versus non-artists is a function of the fact that the vogue in art today is toward the absurd and bizarre? One's personality is affected by his teaching environment and the like. Since artists are exposed in their formal art classes, and in their social environment to the necessity of being bizarre, unrealistic, etc., would this

not be reflected in their preferences for these types of pictures and paintings? Second, if you used an all non-artist population, would not these differences between artist and non-artist disappear? If they did, you would then be unable to judge personality from an analysis of preferences for pictures, paintings or drawings.

DR. MACKINNON: I should like to answer that by saying I would certainly agree with the speaker that the current mode, the cultural values are certainly of the greatest importance in this area. I should, however, be inclined to think that these differences in perceptual choice, the perceptual decision, what it is that individuals tend to prefer to see in their environments, is something which has a long history, and in the forthcoming paper by Frank Barron, Barron makes some extremely interesting comparisons between the typology, if you will, which he has developed on the basis of his investigation in this area and William James' typology of the tough and tender minded, and I think that these come out again in the work which Cyril Burt has done. While I would not minimize the thesis you have stressed, I think there are differences here that go beyond any single period of time.

DR. FLEMMING: I have been very much interested in the tests that have been developed by Dr. Cattell and by Drs. Guilford and Martin which have been a result of factor analysis. Presumably one of the purposes of factor analyzing the responses on paper and pencil tests is to get what might be called a personality structure in terms of unitary characteristics. The psychologists who factor analyze these tests, of course, publish them and sell them to those of us in industry or business who want to use them for the purpose of selection or prediction of behavior. While theoretically, in the tests that have been developed, you have unitary characteristics, in that they are factorially pure, or as pure as they can be gotten, in actual practice I have found, in using the Guilford-Martin tests, that there is some question as to whether or not these factored tests really measure unitary characteristics.

For instance, I have recently made an item analysis of the tests of a thousand applicants for jobs as salesmen, of five hundred applicants for jobs as sales managers and of another three or four hundred tests of men who are actually employed as salesmen and rated satisfactory or better than average by their managers. I suppose the authors of the factored tests, because the factors are statistically pure factors, would say you don't need to make an item analysis. But in making an item analysis of the questions on the Guilford-Martin tests, I find that there are some items which salesmen and sales managers answer in such a way as to indicate that they are thinking extroverts, whereas there are other questions which the majority of them

answer in such a way as to indicate that they are thinking introverts. Specifically, I found that among the questions, which are supposed to measure thinking extroversion, 16 are answered by the *majority* of these salesmen and sales managers in that way. However, 12 of the questions, the *majority* of salesmen and sales managers answer in such a way as to indicate that they are thinking introverts.

Now, a test is of use if it enables us to predict not only the characteristic that the individual has but whether or not, in terms of the score on that characteristic, he will be effective in a given situation. Obviously, with the split of 16 and 12, any score which indicates that a man is a thinking introvert or a thinking extrovert would not predict the effectiveness of salesmen or sales managers.

I find a similar situation with respect to the measure that is called "Agreeableness." There are 11 questions which the *majority* of salesmen answer in such a way as to indicate that they are agreeable; but there are 5 questions that the *majority* of salesmen answer in such a way as to indicate that they are not agreeable. In other words, while statistically these traits may be unitary traits, they seem to me, in practical application, to represent statistical artifacts, which may be misleading to the individual who attempts to make a practical application of the tests and predict the kind of behavior which you would find in a given occupational situation.

DR. CATTELL: So far as I understand that problem, it would mean that the regression equation that is the prediction of goodness of salesmanship is one combining several factors. I would not expect *one* unitary trait to predict performance at all well. The point in my culminating paragraph when I spoke originally was that you are deliberately picking a number of unrelated predictors which only when they are together can give you the maximum prediction. For the rest, that seems to me to deal with questions which Professor Guilford should answer as to the particular reliability of those particular items.

Appendix

PARTICIPANTS—1951 INVITATIONAL CONFERENCE ON TESTING PROBLEMS

ABRAMSKY, Murray, New York City Youth Board
ADAMS, Elizabeth, Educational Testing Service
ADAMS, Joe K., Bryn Mawr College
ADKINS, Dorothy C., University of North Carolina
AHMANN, J. Stanley, Cornell University
ALLEN, Charles D., Educational Testing Service
ALLEN, Margaret E., Maine Public Schools
ALLISON, Jane, University of Chicago
ALLISON, Roger B., Jr., Educational Testing Service
ALMAN, John E., Boston University
ALT, Pauline M., Teachers College of Connecticut
AMOS, Nir, Washington, D. C.
ANASTASI, Anne, Fordham University
ANDERSON, Rose G., Psychological Corporation
ANDERSON, Roy N., North Carolina State College
ANDREWS, T. G., University of Maryland
ANGOFF, William H., Educational Testing Service
APPELL, Morris, Brooklyn College
APTER, Eleanor, Educational Testing Service
ARSENIAN, Seth, Springfield College
ASH, Philip, Pennsylvania State College
AYER, Frederic L., Teachers College Columbia University

BAIER, Donald E., Personnel Research Section, Adjutant General's Office
BANNON, Charles J., Crosby (Conn.) High School
BARRETT, Dorothy M., Hunter College
BARTNIK, Robert V., Educational Testing Service
BECHTOLDT, Harold P., State University of Iowa
BECK, Hubert P., The City College of New York
BECKER, Theodore, New York State Department of Civil Service
BEDELL, Ralph, American University
BEERY, John R., University of Miami
BENSON, Arthur L., Educational Testing Service
BERDIE, Ralph F., University of Minnesota
BERGESEN, B. E., Personnel Press, Incorporated
BITTNER, Reign H., Prudential Insurance Company
BLACKETT, Shirley A., Educational Testing Service
BLACKMER, Alan R., Phillips Academy
BLAUL, R. Elizabeth, Highland Park High School
BLOOM, Ben S., University of Chicago
BOASI, Veronica M., Archdiocesan Vocational Service
BOLLENBACHER, Joan, Cincinnati Public Schools
BRACA, Susan E., Archdiocesan Vocational Serivce
BRAGDON, Henry W., Phillips Exeter Academy
BRANDT, Hyman, American Occupational Therapy Association
BRANDWEIN, Paul F., Teachers College, Columbia University
BRANFORD, Thomas L., New York State Civil Service Department
BRAY, Douglas W., Columbia University
BRIDGES, Claude F., World Book Company
BRISTOW, W. H., Board of Education, New York
BROOKS, Richard B., College of William & Mary
BRYAN, Miriam M., Silver Burdett Company
BRYAN, Ned, Rutgers University
BRYANT, Merrilee M., Educational Testing Service
BUCKINGHAM, Guy E., Allegheny College
BUCKTON, LaVerne, Brooklyn College
BUNKER, Harris F., University of Puerto Rico
BURKE, James M., Storrs (Conn.) Grammar School
BURKE, Paul J., World Book Company
BURKE, William W., Washington University
BURNHAM, Paul S., Yale University
BUROS, Oscar K., Rutgers University
BURT, S. M., Educational Testing Service
BYRNE, Richard Hill, University of Maryland

CAMPBELL, Donald W., Board of Education, Newark, N. J.
CANER, Faruk, The Turkish Embassy
CANFIELD, Frederick, H., Walter Kidde & Company, Incorporated
CAPPS, Marian P., The New York City College
CARLSON, C. R., Maxwell Air Force Base, Alabama
CARLSON, Harold S., Upsala College
CARLUCCI, Cosimo A., Queens College
CARROLL, John B., Harvard University
CARTER, W. R., University of Missouri
CARY, James L., Howard University
CATTELL, R. B., University of Illinois
CAUFIELD, Frederick H., Walter Kidde & Company, Incorporated
CAYNE, Bernard S., Educational Testing Service
CHASIN, Joseph, New York State Civil Service Department
CHAUNCEY, Henry, Educational Testing Service
COBB, Colonel Candler, New York Selective Service System Headquarters
COHEN, Elizabeth L., Educational Testing Service
COHN, Fannia M., I.L.G.W.U., New York, N. Y.
COLEMAN, William, University of Tennessee
COMREY, Andrew L., University of California
CONRAD, Herbert S., U. S. Office of Education, Washington, D. C.
COPELAND, Herman A., Atlantic Refining Co.
CORCORAN, Mary, Educational Testing Service
COWLES, John T., Educational Testing Service
COY, Geneviere L., Dalton School, New York City
CRAIG, Robert, Teachers College, Columbia University
CRANE, Harold L., Jr., Educational Testing Service
CRANE, Percy F., University of Maine
CRISSY, W. J. E., Queens College
CRONBACH, Lee J., University of Illinois
CRUTCHFIELD, Richard S., Swarthmore College
CUTTS, Norma E., New Haven State Teachers College
CYNAMON, Manuel, Brooklyn College

DAILEY, John T., Bureau of Naval Personnel, Washington, D. C.
DALY, Alice T., New York State Education Department
DAVIDSON, Helen H., The City College of New York
DAVIDSON, Hugh M., The Pennsylvania State College
DERRICK, Clarence, Educational Testing Service
DIEDERICH, Gertrude W., Educational Testing Service
DIEDERICH, Paul B., Educational Testing Service
DIERS, Helen A., Teachers College Columbia University
DION, Robert, California Test Bureau
DOLLARD, John, Yale University
DOPPELT, Jerome E., The Psychological Corporation
DRAGOSITZ, Anna M., Educational Testing Service
DRESSEL, Paul L., Michigan State College
DUDEK, Edmund E., University of Washington
DUNN, Frances E., Brown University
DUROST, Walter N., Boston University
DYER, Henry S., Harvard University

EBEL, Robert L., State University of Iowa
EDGERTON, Harold A., Richardson, Bellows, Henry & Co., Incorporated
ELEY, Farle, University of Chicago
EPSTEIN, Bertram, The City College of New York

FAN, Chung-Tch, Educational Testing Service
FAY, Paul J., New York State Department Civil Service
FEDER, D. D., University of Denver
FELS, William C., College Entrance Examination Board
FERGUSON, Leonard W., Aetna Life Insurance Company
FINDLEY, Warren G., Educational Testing Service
FITZPATRICK, G. M., Teachers College of Connecticut
FLANAGAN, John C., American Institute for Research
FLEMING, Mae E., Educational Testing Service
FLEMMING, Cecile White, Burton Bigelow Organization
FLEMMING, Edwin G., Burton Bigelow Organization
FOLEY, John P., Jr., The Psychological Corporation
FRANKLIN, Jean, Educational Testing Service
FREDERIKSEN, Norman, Educational Testing Service
FREEMAN, Paul M., Maxwell Air Force Base, Alabama

FRENCH, John W., Educational Testing Service
FRUCHTA, B., University of Texas
FURST, Edward J., University of Michigan

GAGE, Barbara, Educational Testing Service
GALLAGHER, Henrietta L., Educational Testing Service
GALSTON, Samuel H., Municipal Civil Service Commission
GARDNER, Eric F., Syracuse University
GARRISON, Harry, Educational Testing Service
GERBERICH, J. Raymond, University of Connecticut
GIANGRANDE, Salvatore C., St. John's University
GIBBONS, Mary L., Catholic Charities, New York
GILMORE, John V., Boston University
GOODMAN, Samuel M., Maxwell Air Force Base, Alabama
GRANT, Donald L., Prudential Insurance Company
GREEN, Dorothy, Bolling Air Force Base, Washington, D. C.
GREEN, Joseph C., Department of State, Washington, D. C.
GREENE, Edward B., Wayne University
GRIMM, Elaine R., American Public Health Association
GULLIKSEN, Harold, Educational Testing Service
GUSTAD, John W., Vanderbilt University

HADDAD, R. K. New York University
HAGEN, Elizabeth, Teachers College Columbia University
HAGMAN, E. R., Greenwich Public Schools
HASTINGS, J. T., University of Illinois
HAVIGHURST, Robert J., University of Chicago
HEDLUND, Paul A., New York State Educational Department
HICKS, Allen, Ministry of Education, London
HIERONYMUS, A. N., State University of Iowa
HILL, Walker H., Michigan State College
HOBERMAN, Solomon, Municipal Civil Service Commission, New York
HOGAN, Ralph M., Office of Naval Research
HOLLINSHEAD, Byron, Commission on Financing Higher Education
HOROWITZ, Milton W., Queens College
HUDDLESTON, Edith M., Educational Testing Service
HUMPHREYS, Lloyd G., Lackland Air Force Base, San Antonio, Texas
HUNT, Barbara S., Educational Testing Service
HURD, A. W., Medical College of Virginia
JACOBS, Robert, A. & M. College of Texas
JASPEN, Nathan, Pennsylvania State College
JENNINGS, Helen H., Brooklyn College
JOHNSON, A. Pemberton, Educational Testing Service
JOHNSTON, John R., U. S. Civil Service Commission
JONES, Edward S., University of Buffalo
KABACK, Goldie R., The City College of New York
KARLIN, John E., Bell Telephone Laboratories
KAVENEY, Ethel, Educational Testing Service
KELLEY, DeCourcy, Educational Testing Service
KELLEY, H. Paul, Educational Testing Service
KERNAN, John P., Dunlap and Association, Incorporated
KLARE, George R., The Psychological Corporation
KOGAN, Leonard S., Institute of Welfare Research
KRUGLOV, Lorraine, The City College of New York
KUDER, G. Frederic, Duke University
KURTZ, Albert K., Pennsylvania State College

LANGMUIR, Charles R., Syracuse University
LANNHOLM, G. V., Educational Testing Service
LAYTON, Wilbur L., University of Minnesota
LEACH, Kent W., University of Michigan
LEDIG, H. M., Institut für Schulpsycholog., Berlin, Germany
LENNON, Roger T., World Book Company
LEVERETT, Hollis M., American Optical Company
LLOYD, Dorothy A., Educational Testing Service
LONG, Louis, The City College of New York
LORD, Frederic M., Educational Testing Service
LORGE, Irving, Teachers College Columbia University
LORR, Maurice, Veterans Administration, Washington, D. C.
LUCAS, Charles M., Educational Testing Service
LURIE, Walter A., National Community Relations Advisory Council

MACKINNON, Donald W., University of California

112

MacPhail, Andrew H., Brown University
McArthur, Charles C., Harvard University
McCall, W. C., University of South Carolina
McCambridge, Barbara, Educational Testing Service
McCollum, Joyce E., New York Department of Civil Service
McCord, Mary, University of Chicago
McCully, C. Harold, Veterans Administration, Washington, D. C.
McGuire, M. Anne, New York State Department of Civil Service
McIntire, Paul H., University of New Hampshire
McKinney, Lidie, Educational Testing Service
McGillicuddy, Marjorie, New York State Department of Civil Service
McLaughlin, Kenneth, U. S. Naval Academy
McQuitty, John V., University of Florida
Manuel, H. T., University of Texas
Marsh, Donald D., Pace College
Martin, Lycia O., State Teachers College, Trenton, N. J.
Masia, Bertram B., American Public Health Association
Mayans, Frank, Teachers College, Brooklyn, N. Y.
Mays, Luzelle D., U. S. Civil Service Commission
Mazza, Marguerite, Greenwich Public Schools
Messick, Samuel J., Educational Testing Service
Mendleson, Martin A., Test Unit No. 3, U. S. Air Force
Michael, William B., Rand Corporation, Santa Monica, California
Mieler, Peter M., Educational Testing Service
Miller, Raymond W., F. A. O. United Nations
Mitchell, Blythe, World Book Company
Mollenkopf, William G., Educational Testing Service
Morrison, Thomas F., Milton Academy
Moses, Lincoln E., Teachers College, Columbia University
Muhyi, I. A., Teachers College, Columbia University

Nielander, William A., Hofstra College
Nixon, W. D., Teachers College, Columbia University
North, Robert D., University of Kentucky

Olsen, Marjorie, Educational Testing Service
Orleans, Beatrice, U. S. Navy, Washington, D. C.
Orleans, Isaak D., Mitchell Air Force Base, New York
Orleans, Joseph B., George Washington High School
Oxtoby, Toby, Commission on Human Resources and Advanced Training

Partington, Dorothy B., Educational Testing Service
Pashalian, Siroon, Fordham University
Perlman, Mildred, New York Civil Service Commission
Perloff, Robert, Research Section, Department of the Army, Washington, D. C.
Perry, W. D., University of North Carolina
Peterson, Donald A., Life Insurance Agency Management Association
Peterson, William C., Standard Oil Co. of New Jersey
Phillips, Laura M., Silver Burdett Company
Plumlee, Lynnette B., Educational Testing Service
Polin, A. Terrence, Teachers College, Columbia University
Poteete, Robert A., *New York Herald Tribune*
Potts, Edith Margaret, The Psychological Corporation
Powers, James T., U. S. Civil Service Commission
Prescott, George A., World Book Company
Preston, Braxton, Educational Testing Service

Quick, Robert, American Council on Education

Ramirez, Jesus M. Colon, Teachers College, Columbia University
Raskin, Evelyn, Brooklyn College
Reid, John W., Teachers College, Columbia University
Remmers, H. H., Purdue University
Rhule, Warren Allen, Educational Testing Service
Ricciuti, H. N., Educational Testing Service
Ricks, J. H., The Psychological Corporation
Rimalover, Jack K., Educational Testing Service
Rivlin, Harry N., Queens College
Roca, Pablo, Department of Education, Puerto Rico
Rock, R. T., Jr., Fordham University
Rogers, Charlotte Fox, Educational Testing Service

ROGERS, Miles S., Educational Testing Service
ROWDEN, Dorothy, Markle Foundation
RULON, Phillip J., Harvard University
RUSSELL, James T., New York State Department of Civil Service

SAUNDERS, David R., Educational Testing Service
SCATES, Douglas E., Queens College
SCHRADER, William B., Educational Testing Service
SCHROEDEL, E. C., International Business Machines Corporation
SCHULTZ, Douglas, Educational Testing Service
SCHULTZ, Margaret K., Educational Testing Service
SCHWEIKER, Robert, Educational Research Corporation
SEASHORE, Harold, The Psychological Corporation
SEIDENFELD, Morton A., National Foundation for Infantile Paralysis
SHARP, Catherine G., Educational Testing Service
SHAYCOFT, Marion F., American Institute for Research
SHERBURNE, J. W., Oregon State College
SHIELDS, William S., U. S. Naval Academy
SHIENBLOOM, Charles, Board of Public Education
SIDES, Robert W., Phillips Academy
SMITH, Alexander F., University of Connecticut
SMITH, Allan B., University of Connecticut
SPANEY, Emma, Queens College
SPAULDING, Geraldine, Educational Records Bureau
SPEER, George S., Illinois Institute of Technology
SPENCE, Douglas, U. S. Military Academy
SPENCE, Ralph B., Teachers College, Columbia University
SPEROFF, B. J., Psychometric Affiliates, Chicago, Illinois
STALLMAN, F. B., New York Telephone Company
STANLEY, Julian C., George Peabody College for Teachers
STERNBERG, Jack J., Harvard University
STEWART, Naomi, Educational Testing Service
STONE, Paul T., Huntingdon College
STOUFFER, S. A., Harvard University
STOUGHTON, Robert W., Connecticut State Department of Education
STURM, Joseph, New Lincoln School, New York City
SULLIVAN, Richard H., Educational Testing Service
SUPER, Donald E., Teachers College, Columbia University
SYMONDS, Percival M., Teachers College, Columbia University
SWANSON, Edward O., Educational Testing Service
SWIFT, Everett L., The Peddie School
SWINEFORD, Frances, Educational Testing Service

TABER, Victor A., New York State Education Department
TASSO, Charles A., Richardson, Bellows, Henry & Company
TAYLOR, Calvin W., University of Utah
TAYLOR, Donald W., Stanford University
TAYLOR, Justine N., Educational Testing Service
TCHORNI, Jacquelyn G., Educational Testing Service
TCHORNI, Bernard L., Educational Testing Service
TERRAL, Joseph E., Educational Testing Service
THORNDIKE, Robert L., Teachers College, Columbia University
TINKLEMAN, Sherman, New York State Education Department
TINKLE, J. W., Mitchell Air Force Base, Long Island, N. Y.
TODD, Charles B., Mitchell Air Force Base, Long Island, N. Y.
TOMKINS, Silvan, Princeton University
TRAVERS, Robert M. W., Board of Higher Education, New York
TRAXLER, Arthur E., Educational Records Bureau
TREMER, Carl E., New York State Civil Service Department
TRIGGS, Frances, Committee on Diagnostic Reading Tests, Incorporated
TUCKER, Ledyard R, Educational Testing Service
TURNBULL, William W., Educational Testing Service
TYLER, Matilda, Yale University

UPSHALL, Charles C., Eastman Kodak Company

VOTAW, D. F., Jr., Yale University
VREDEVOE, L. E., University of Michigan

WALKER, Helen M., Teachers College, Columbia University
WATSON, Walter S., The Cooper Union
WECHSLER, David, New York University
WENZEL, Bernice M., Barnard College

WERENFELS, Eleanor H., Educational Testing Service
WESMAN, A. G., The Psychological Corporation
WHITMORE, Howard S., New York Telephone Company
WILKE, Walter H., New York University
WILKS, S. S., Educational Testing Service
WILLIAMS, Malcolm J., U. S. Coast Guard Academy
WINANS, S. D., New Jersey State Department of Education
WINGO, Alfred L., Virginia State Department of Education
WINKLER, Lila, New York University
WODELL, Blandena, World Book Company
WOLZ, Charles, New York State Department of Civil Service
WOOD, Ray G., Ohio State Department of Education
WORTMAN, Richard E., Educational Testing Service
WRIGHT, Wilbur H., New York State University Teachers College
WRIGHTSTONE, J. Wayne, Board of Education of the City of New York
YEOMANS, Alice V., American Council on Education
YORUKOGLU, Kadri, Council of Education, Turkey
ZIFF, Larzer, Educational Testing Service
ZUBIN, Joseph, Columbia University

150